DOCUMENT RESUME

ED 097 382 88 TM 004 131

AUTHOR Shaffer, J. Stephen, Jr.
TITLE An Assessment of the Title III, ESEA Validation Effort, 1973-74. Final Report.
INSTITUTION Scientific Management Associates, Gloucester, N.J. Educational Systems Div.
SPONS AGENCY Bureau of Elementary and Secondary Education (DHEW/OE), Washington, D.C.
PUB DATE 74
CONTRACT OEC-0-74-1412
NOTE 129p.; For a related document, see ED 081 851

EDRS PRICE MF-$0.75 HC-$6.60 PLUS POSTAGE
DESCRIPTORS *Evaluation Methods; *Program Evaluation; Project Training Methods; Self Evaluation; Team Training; Training Techniques; Validity; Workshops
IDENTIFIERS *Elementary Secondary Education Act Title III; ESEA Title III

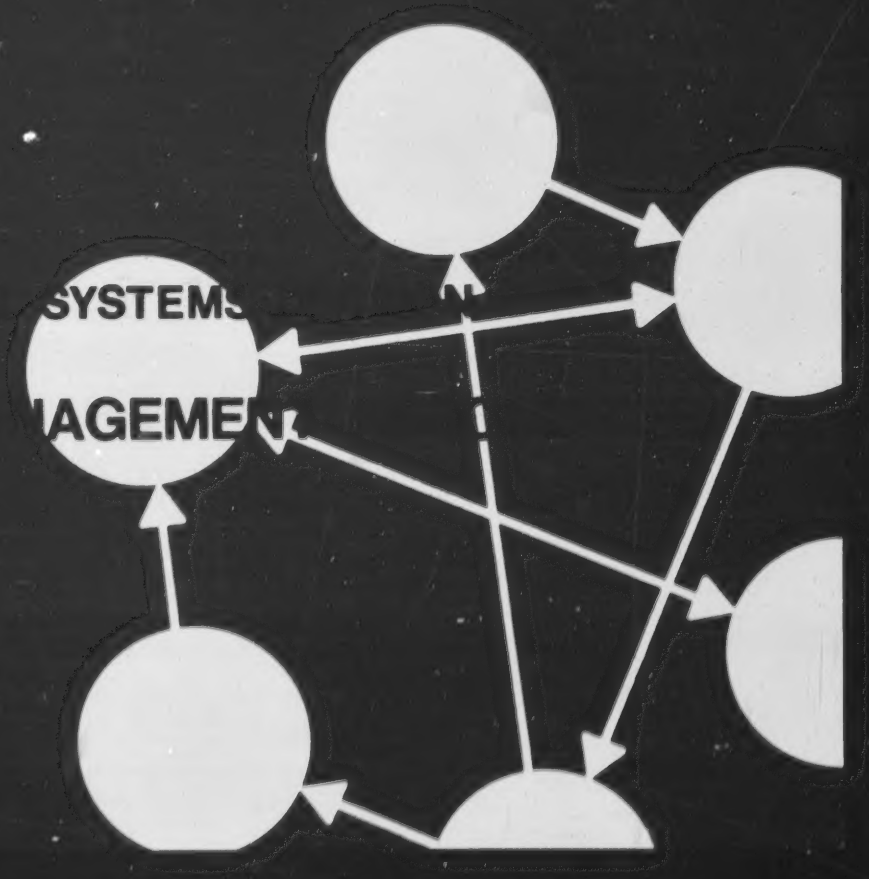
ABSTRACT

Evaluation materials and instructional content relevant to the IVD process for 1973-74 are presented. Section 1 of this report is an analysis of on-site experience of the validators in the employment of the Validator Self-Analysis Forms. Section 2 is Scientific Management Associates Educational Systems Division (SMA/ESD) evaluation of the 10 national training workshops for validators and state and local project personnel. Section 3 includes the authors' recommendations for year 3 developmental activities based on input from the validators' critiques, the participant evaluation of workshops, and SMA/ESD's involvement in the instrument design and training phases. The appendices contain a prospectus, the content for validator training workshops, and a partial list of candidate projects visited, 1974. (Author/RC)

SMA

ED 097382

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY



FINAL REPORT

AN ASSESSMENT OF THE TITLE III, ESEA VALIDATION EFFORT, 1973-74

TABLE OF CONTENTS

	Page
Preface	i, ii
Part I	
The Validation Effort as Assessed by the Validation Teams: A Content Analysis of the Validator Self-Analysis Form, 1973-74	1
Part II	
The Interim Report on the Conduct of the Validator Training Workshops	17
Part III	
Recommendations	32
Part IV	
Appendices	
Appendix A - Prospectus	
Appendix B - Content for Validator Training Workshops	
Appendix C - Partial List of Candidate Projects Visited, 1974	

Preface

This Final Report document represents an effort on the part of the Educational Systems Division of Scientific Management Associates to gather together under one cover the evaluation materials and the instructional content pertinent to an understanding of the IVD process for 1973-74. A similar document was prepared last year by virtue of SMA/ESD's involvement in the IVD process and was accepted by the ERIC file. Its reference appears in the January, 1974 issue of ERIC, ED #081 851.

Section I of this report is an analysis of the on-site experience of the validators in the employment of the Instrument. Section II is an SMA/ESD evaluation of the ten national training workshops for validators, state and local project personnel -- heretofore submitted as an Interim Report to USOE/DSCS in May, 1974. Section III includes our recommendations for Year III developmental activities based on input from the validators' critiques, the participant evaluation of workshops, and SMA/ESD's involvement in the Instrument design and training phases.

A special note of thanks is directed to Dr. Lee Wickline and his entire staff without whose unselfish assistance would have made this second year's IVD training effort difficult at best: to Mr. Gerald Kluempke, Executive Secretary of the Title III National Advisory Council; to all of the State Title III coordinators for their organizational efforts and especially to those ten regional host coordinators who

planned the training sessions; to my secretary, Anna Martin, for her devotion; and last, but certainly not least, to all of those workshop participants -- validators, project personnel, State Title III staffs -- who supported the training sessions with their excellent contributions in spite of the exigencies of time and facilities.

J. Stephen Shaffer, Jr.

Gloucester, New Jersey
July 8, 1974

PART I

A Content Analysis of the Validator Self-Analysis Forms

A content and inter-item analysis of the Validator's Self-Analysis (on-site) Form[1] indicates positive steps for the improvement of the I.V.D. procedure, and for the extensive and much needed revision of the validation instrument itself. It should be noted, however, that it was generally echoed throughout all ten regional training sessions that this year's I.V.D. Instrument ('73-'74) was a vast improvement over last year's.

There were 177 Self-Analysis forms returned representing a clear majority of those serving on validation teams. An exact count is not possible for we do not have figures on validators serving on more than one team, nor the total number of on-site visits completed.[2] It is our opinion, however, that the number of responses provides a thoroughly sufficient basis upon which to make generalizations.

The breakdown by Section of the Instrument is as follows:[3]

	No.	%
Effectiveness/Success	75	43
Cost Information	45	25
Exportability	57	32
Totals	177	100

---

[1] Please see next page.

[2] Please see Appendix C. This is a list of the projects visited from which the 177 Self-Analysis forms were completed and returned and from which the data herein were tabulated, analyzed and interpreted.

[3] The numbers entered in this chart by Section are not to indicated that Section exclusively. Some respondents functioned in all three Sections. Based on those respondents' answers to the other questions on the Form, this writer arbitrarily assigned them to a specific Section.

Validator Self-Analysis Form

Please fill in the following information at the conclusion of the on-site validation and mail in the envelope enclosed with the Guidebook.

Name_____ Date_____

Project Reviewed for Validation_____

City_____ State_____

Your Address_____

City_____ State_____

Section Reviewed_____

1. Were you a validator last year? Yes_____ No____

2. If yes, did you validate the same criterion as last year? Yes_____ No____

3. Do you feel your involvement reflected your area of expertise? Yes_____ No____

4. Were the point values for the section you validated generally acceptable to your teammates? Yes_____ No____

5. Did you find the task of assessing the data and weighing the responses difficult? Yes_____ No____

   Comments:_____

   _____

   _____

6. Is this year's validation instrument generally better or worse than last year's? Better____ Worse____

7. Which section of the instrument was most improved over last year's?

   Effectiveness/Success_____ Cost Information_____ Exportability____

8. Were there questions in your section that you found particularly difficult to answer in terms of assessing number weights? Yes___ No__
   If yes, please indicate section and question numbers

   _____

   _____

9. Do you feel that there was adequate team interaction and discussion in reaching a conclusion on each of the three sections of the report?

   Yes_____ No____

10. Please comment on areas of difficulty with respect to both the validation instrument, and the team's interaction with one another and with project personnel.

    _____

    _____

Questions 1, 2, 3, 4, 6, 7 and 9 of the Self-Analysis Form are essentially closed questions (all dichotomous choices except #7) and their breakdown, question-by-question, is as follows:[4]

Question 1: "Were you a validator last year?"

Yes	%	No	%
70	40	107	60

Question 2: "Did you validate the same criterion as last year?"

Yes	%	No	%
44	63	26	37

Question 3: "Do you feel your involvement reflected your area of expertise?"

Yes	%	No	%
168	97	5	3

Question 4: "Were the point values for the section you validated generally acceptable to your teammates?"

Yes	%	No	%
170	100	0	0

Question 6: "Is this year's validation instrument generally better or worse than last year's?"

Better	%	Worse	%
63	84	12	16

---

[4] The reader will please note that totals for certain questions do not correspond with the total number of respondents. This is due to the fact that a negligble number of repondents failed to answer the question(s) because of either inadvertance or oversight.

Question 7: "Which section of the instrument was most improved over last year's?"

	No.	%
Effectiveness/Success	38	49
Cost Information	22	28
Exportability	18	23
Total	78	100

Question 9: "Do you feel that there was adequate team interaction and discussion in reaching a conclusion on each of the three sections of the report?"

Yes	%	No	%
171	99	1	1

Of those responding 107 or 60% were new validators. This amount was suprisingly high given last year's effort to establish a validator "bank". In spite of the reasonably large number of new validators, the close correspondence in totals for questions 3, 4 and 9 suggests that one may conclude that the validator selection procedure and the trainer's instruction for team interaction were generally productive, and can be recommended for continuation in Year III development activities.

Similarly, there is close correspondence in the totals for questions 6 and 7 -- 75 and 78 respectively. It remains curious to this writer why so many of the respondents -- around 100 -- chose not to answer these questions. The revised Instrument elicited the full range of emotions in the training sessions -- from those who felt that it was a quantum improvement over last year's on one end, to those who felt it was less rigorous. Lacking a sufficient number of responses to questions 6 and 7, this writer cannot qualify a conclusive statement about the improved

quality of this year's Instrument. However, from purely subjective observation and informal conversations with many involved people, it is this writer's considered judgement that the 1973-74 IVD Instrument is a substantial improvement over last year's.

Of those responding to question 5 of the Self-Analysis Form, 80 or 47% found the task of assessing the data and weighing responses difficult, whereas 90 or 53% indicated either little or no difficulty (or "difficult but not impossible") in responding.[5] Similarly, in question 8, 42 or 27% of those responding indicated they experienced difficulty in responding to questions in their assigned sections. For purposes of comparison, the responses to these questions, 5 and 8, by class were as follows:

Ouestion 5: "Did you find the task of assessing the data and weighing the responses difficult?"

	Yes	%	No	%
Effectiveness/Success N=73	37	51	36	49
Cost Information N=44	15	34	29	66
Exportability N=53	27	51	26	49

---

[5] As another form of analysis for Year III, it would be appropriate to compare validators' responses on the basis of projects they worked on that were approved or disapproved. Correlations of responses would assist in identifying where non-validated projects fall down in the documentation proceedings.

These responses overall indicate that 45% of the respondents had difficulty in answering questions in their assigned section. The above chart shows a 17 point range for all three classes. Answering the questions in the Instrument required a two step procedure including 1) assessing the adequacy, appropriateness and accuracy of the data (documentation and testimony), and 2) assigning number weights. Since question 5 of the Form is a two part question, it is not possible to identify the greater area of difficulty in responding. It is this writer's position, however, that since the two judgements are inextricably related, the feedback indicates the need for another major Instrument revision responsive to the problematic questions identified in response to question 8 of the Self-Analysis Form.

Responses to Question 8 were as follows:

"Were there questions in your section that you found particularly difficult to answer in terms of assessing number weights"

	Yes	%	No	%
Effectiveness/Success N=63	18	29	45	71
Cost Information N=44	11	25	33	75
Exportability N=52	13	25	39	75

The responses overall indicate that 26% of the respondents had a problem in answering particular questions in their assigned sections. This time the response range among the three classes is only four points. This is a significant reduction in problematic questions from last year's IVD effort. It remains clear, however, that those serving as validators in Sections I and III -- Effectiveness/Success and

Exportability -- experienced the greatest difficulty in responding.

It is our recommendation that revision of the Instrument be continued on all Sections with particular attention to Sections I and III as the data indicate.

The following chart shows percentage responses by classes for questions 3, 4, 5, 8 and 9, all of which have inter-relationship in this item analysis with particular bearing upon the team interaction component of the IVD process.

Question 3: "Do you feel your involvement reflected your area of expertise?"

Question 4: "Were the point values for the section you validated generally acceptable to your teammates?"

Question 5: "Did you find the task of assessing the data and weighing the responses difficult?"

Question 8: "Were there questions in your section that you found particularly difficult to answer in terms of assessing number weights?

Question 9: "Do you feel that there was adequate team interaction in reaching a conclusion on each of the three sections of the report?"

Percentage Responses to Self-Analysis Form

		Ques. 3	Ques. 4	Ques. 5	Ques. 8	Ques. 9
I Effectiveness/Success N=71	Yes	100%	100%	51%	29%	100%
	No	0	0	49	71	0
II Cost Information N=44	Yes	98%	100%	34%	25%	100%
	No	2	0	66	75	0

		Ques. 3	Ques. 4	Ques. 5	Ques. 8	Ques. 9
III Exportability N=55						
	Yes	99%	100%	51%	25%	99%
	No	1	0	49	75	1

The above analysis clearly indicates that the areas of difficulty are those of assessing the adequacy, accuracy and appropriateness of the data. This is compounded by the lack of parameters (or operational definitions) of the terms, "persuasive", "substantial", "conclusive" and "compelling".

General Areas of Difficulty

The following list of areas of difficulty is complied from responses to questions 5 and 10 on the Self-Analysis Form. It is ordered from the most to the least frequency of occurrence.

Validation Instrument

* 1. not enough time to do the job
* 2. validation instrument insufficient/ambiguous
* 3. too complex
* 4. project information was minimal/inconsistent
* 5. too many overlapping areas/too much repetition and duplication
* 6. far too many Instrument pages/too much subjectivity
* 7. state departments should have reviewed data more carefully
* 8. many revisions necessary
* 9. LEA's should have validation requirements upon initiating projects
* 10. too research oriented/too much emphasis on statistics
* 11. too much additional data needed/project applications incomplete
* 12. statement of objectives needs more screening at SEA level

---

* categories into which the majority of remarks fell

* 13. Instrument should be expanded to get at relevant variables/more open-ended types of questions

* 14. the inequitable distribution of asterisked questions per section, i.e., an entire project could fall on not meeting the 3-point maximum on question 4(a) in Section I: or by not meeting the 3-point maximum on one question in Section II.

* 15. the inapplicability of certain questions

* 16. the issue of project procedures or process objectives not meeting the "hard" data requirements when in many cases the project's contribution to innovativeness/exemplariness was the process utilized

* 17. the Instrument was too dependent upon well stated objectives, but without reference to the meaningfulness of the objectives

* 18. the Instrument seems to demand a research design methodology and is, therefore, unresponsive to certain types of affective- and psychomotor- centered projects

* 19. the various terms in the Instrument are inadequately defined

* 20. the need for a clearer distinction between the role of the validator and the evaluator, i.e., some few questions demanded not the review of documentation but the assessment of the value of the particular practice or procedure

Team Interaction

* excellent

* worked well together

* exceptionally cooperative and diligent workers

Team Interaction with Project Personnel

* excellent

* should have been better organized for on-site visitation

* interaction very poor

Miscellaneous Comments

* insufficient time to review projects
* training sessions need to be improved
* too much paper work
* late availability of informational material for validation

Specific Areas of Difficulty: Section I

* multi-faceted questions, i.e., answering questions with multiple, sometimes conflicting reference points. E.g., the problem of "educational" and "statistical" significance
* the issue of operational definitions for "accuracy", "appropriateness", "substantial", etc., i.e., what do these terms mean?
* the questions of "validity" and "reliability" for project generated tests
* the issue of poorly stated objectives; the on-site revision of objectives, the absence of objectives, unrealistic performance levels
* the issue of erratic documentation
* the issue of validator role confusion, e.g., "is the function evaluation, validation or auditing?"
* the issue of inadequate evaluation designs well completed (with resultant high scores) versus a more adequate design possibly less satisfactorily completed, i.e., is the Instrument too research oriented?
* the issue of projects with incomplete data either because of mid-point development, inadequate orientation to the Instrument's data requirements, or generally inadequate or inconclusive documentation
* the issue of hearing evidence through on-site interviews and building such evidence into responses to Instrument questions

Questions Causing Difficulty in Section I

The questions mentioned as causing Section I validators the greatest amount of difficulty are as follows and in this order of priority:

1. Question 4 - The entire subset of questions (4a - i) pertaining to the evaluation design of the project.

2. Question 7 - "Review and verify the evidence supporting the conclusions that the findings for the nominated objective and determine the adequacy of the evidence of need to justify the selection of the objective."

3. Question 2 - "Examine and verify the needs assessment procedures and findings for the nominated objective and determine the adequacy of the evidence of need to justify the selection of the objective."

   Question 1 - "Review the structure of the objectives . . .: (1) who is able to do what, (2) at what level of performance, and (3) under what conditions. . . ."

4. Question 3 - "Examine and verify the activities (methods, strategies, program intervention, etc.) employed to accomplish the objective. Verify the intensity of each method in terms of full-time equivalent professional and non-professional personnel required, hours of instruction, etc."

Specific Areas of Difficulty: Section II

Specific areas of difficulty for validators responding to Section II (Cost Information) include many of those previously stated. Essentially, just two concerns surfaced from the Self-Analysis Forms about this section: (1) There was virtually no scale of judgement to be made for any of the five questions in this section, even though a modified Likert-type scale was provided. Each question required a dichotomous choice -- and an entire project could stand or fall on just one question from this section; and (2) that data confined to last year's costs could be misleading. There was general validator

recognition that while the questions dealt with the "costs", they did not deal with "benefits", and that the data collection forms, while a great improvement over last year's, were still inadequate -- especially in terms of trying to establish a cost for an objective, which is what Section I inherently demanded but which Section II -- nor Section III -- met in the formulation of questions.

There was no particular question(s) which caused more problems than others. It is our recommendation that questions for this section be revised to provide opportunities to make comparisons among costs, achievement and benefits -- and thereby, possibly cost-effectiveness.

Specific Areas of Difficulty: Section III

Specific areas of difficulty for those responding to Section III (Exportability) are as follows:

* the difficulty in assessing the potential adopter variables
* the absence of accurate descriptions of institutional variables, e.g., "home" and "community", "school administration", "teaching staff", etc.
* the general absence of documentation (materials) responsive to replication
* the near impossibility of assessing for number weighting the extent of community involvement and/or support
* the absence of data relative to the need for a staff with "special qualifications"

Questions Causing Difficulty in Section III

1. Question 6: "Examine and verify the descriptions of any community and home variables, e.g., the necessity for parental and community involvement."

2. Question 9: "Examine and verify the descriptions of the types, numbers and special qualifications of personnel required for the project."

   Question 10: "Examine and verify the descriptions of the procedures and materials necessary for personnel training."

   Question 11: "Examine and verify the claim that the project can be adopted in whole or in part."

3. Question 7: "Examine and verify the description of the activities determined by the project staff to be critical to the success of the project."

4. Question 2: "Examine and verify the evidence that the project will be continued with State or local funds after the termination of Federal funds."

   Question 5: "Examine and verify the descriptions of institutional variables critical to the success of the project, i.e., school administration, teaching staff, physical facilities."

SMA/ESD is in understandable agreement agreement with the identification of the areas of difficulty (See Section: General Areas of Difficulty, p. 8). The validators, however, have identified both problems within the Instrument, as well as difficulties with philosophy of the procedure itself. We have called out the "areas of difficulty" those eight areas dealing with the questions themselves, as well as areas we believe are training problems, and the remaining number are problems of a philosophical and procedural nature. The statements are essentially the same as we made last year with only minor additions and deletions.

- problems with the Instrument per se
  - a) the limitation of the Instrument
  - b) the Instrument's dependency on well stated objectives
  - c) the need for clarification and redefinition of terms

d) the need to revise the point values to reflect the effectiveness/exportability of the questions

e) the resultant need (of changed point values) to revise the conversion tables

f) the inapplicability of several of the questions (as well as those questions requiring program evaluation rather than the validation of documentation)

g) the need to remove multifaceted questions and to replace them with singly focused questions

h) the need to distribute more equitably the questions by section

. problems requiring emphasis in validator training

a) assist state coordinators and local personnel in the preparation of project documentation

b) resolution of the issue of how behavioral objectives are to be stated (for the purpose of the IVD procedure and the areas the objectives must address, i.e., "domain" objectives relative to instructional practices, and "process" relative to program management)

c) The problem of providing sufficient time to prepare an adequate training program based on simulated validation experiences drawn from the two years of the IVD effort. And that these training experiences be uniformly conducted to preserve the high probability of coder reliability in the gathering of data.

d) the choice of facilities for training be re-examined so as to be more conducive to the task of absorbing a great deal of information in a very short period of time

· procedural/philosophical issues

a) the IVD process presently assumed in the Handbook requires a research methodology with heavy emphasis on some type of experimental design -- pre/post testing

b) There is no distinction in the Instrument between questions requiring the validator to make program evaluations and validation assessment. This confusion of roles biases the validator's response toward evaluating the program beyond what is presented in the documentation, e.g., the relevance of the stated objectives, assessing project information not germane to the objectives cited, and the conscious (or unconscious) desire to evaluate project management procedures even though not included in the documentation.

c) The need to require project data on management procedures such that projects with primary contributions to management can be recognized, and/or projects with effective practices as a result of good management can be seen and validated in their entirety.

d) The question of how to correctly define the parameters of a successful project, i.e., "Does the validator respond only to what can be documented or is he responsible for making a separate determination of the project's gestalt?"

It occurs to us that these four procedural/philosophical issues accentuate the dilemma of defining the validator's role. If the validator is to make program assessments (beyond written documentation submitted), then the process can never be reliable in the sense that all validators are responding to data in a uniform way. If, on the other hand, the validator is to respond solely on the basis of written documentation, or on testimony received from local sources, then the projects will suffer until such time that educators have become skilled documentarians and validators have thrown off their evaluation-oriented biases. Neither possibility appears likely in the near future. There is, nonetheless, considerable cause for optimism inasmuch as the IVD procedure is underway and there is detailed feedback and a bank of experienced personnel going into Year III development activities.

Clearly the next step is the revision of the Instrument vis a vis these multiple criticisms.

Additionally, as a developmental effort, there need be no demand that the Instrument "stand alone" as a validated document until educators on all levels have had more experience in identifying clearly the factors essential for success. In our opinion it is not desirable to insist that the Instrument be required to stand alone as if it were a nationally normed and validated procedure. The issues confronted in validating success on a cost effective replication basis are not unlike the practice of the law. The law does not stand without interpretation, and the entire legal procedure is constantly in a state of development. The gap that needs to be filled between the profession of a law and the profession of certified educational practices is that of developing quantification procedures of general acceptability throughout the total educational community. The IVD Handbook is a first step in this direction.

Part II INTERIM REPORT

Validation Training for Title III, ESEA, Practices

I. Introduction

SMA/ESD personnel conducted ten validation training sessions and one orientation session for USOE/DSCS personnel during the month of March 1974.

The USOE/DSCS orientation session was held in Washington, D.C. and the ten regional meetings were held in:

1. Nashville, Tenn.
2. Raleigh, N.C.
3. Wakefield, Mass. (Boston)
4. Council Bluffs, Ia.
5. Milwaukee, Wisc.
6. Seattle, Wash.
7. Richardson, Tex. (Dallas)
8. Washington, D.C.
9. Phoenix, Ariz.
10. Pittsburgh, Penna.

SMA/ESD personnel were responsible for a workshop of 1½ days duration which followed, generally, a day's orientation to validation conducted by the Regional Coordinator.

Our presentation covered the following areas:

1. An overview of Title III in terms of planning and development techniques utilized by one SEA. This was used to provide a perspective of Title III for those entirely new to the validation process.
2. A theoretical introduction to the task of validation as presented in the document "Sharing Educational Successes: A Handbook for Validation of Educational Practices," February, 1974.
3. The mechanical/logistical details of the validators being on-site, and conducting their documentation review and assessment.
4. A detailed examination of each question in the instrument, as well as a thorough review of pertinent material not covered in the Handbook e.g., differentiating between the roles of evaluator and validator.

Additionally, SMA/ESD staff prepared and distributed a "Guide for On-Site Validation Team Procedures for Title III, E.S.E.A. Practices." Over 650 copies of this 71 page document were distributed. Included in the Guide was a Validator Self-Analysis Form (p. 69) which was to be completed by each validator and returned to USOE in a return addressed envelope provided by the contractor.

As part of the training session, workshop participants were administered pre- and post-tests. (The pre-test was an integral part of the Guide.)[1]

---

1. Please note Guide attached, pgs. 1-4. A copy of the post-test is also attached.

A series of prepared acetate transparencies were utilized on an overhead projector to introduce workshop participants to the validation procedures and theory. Another set of hand-prepared transparencies were prepared "on-the-run" as a result of constructive feedback from all of the training sessions. While the instructional content remained constant throughout the ten regional workshops, the method of presentation was modified and expedited as a result of the preparation of the transparencies, and the resolution of previously unanswered questions prompted by the participants.

A following section provides a thorough analysis of workshop participant responses to the SMA/ESD presentation, and indicates a generally affirmative reaction to content and presentation modes. These findings are gratifying particularly in light of the unavoidable reliance on the spoken word as the primary means of communication.

There was unquestionably an abundance of information to be covered in a short time period. While many of the concepts of validation were carried over from last year's effort, there were, nevertheless, many new concepts that were demanding intellectually and required considerable discussion for full clarification. Pre/post-test results indicate a sufficient grasp of all new and old material. (see Fig. 4).

Finally, we were pleased with the general tones and conduct of the workshops. There was some objection to the procedures used, but upon explanation of the "givenness" of the instrument, the exigencies of time and the demanding training schedule, the pace and the atmosphere were progressively accelerated and cordial. Our recommendations for Year III

training would be that:

a. more time be allotted for a training package to be developed

b. project personnel whose practices are to be validated should not attend the workshops

c. team assignments and team leaders should be determined prior to the training session

d. regional procedures should be standardized

We received extensive feedback on each of the Instrument's questions. We anticipate the opportunity of participating in a continuing relationship with the Division of Supplementary Centers and Services in order to assist in the revision of the instrument in light of the many excellent criticisms. Additionally, we at SMA/ESD have a number of structural recommendations to make relative to the field-use of the Instrument and Handbook, and for validation procedures for the Year III revision.

II. Workshop Participant Responses

Approximately 41% of those participating in the workshops responded to the Post Training Session Reaction Form.[2] This percentage represents 267 responses out of approximately 650 participants. This latter figure was determined on the basis of the figures submitted to us by the Regional Title III Coordinators. Those responding to the reaction form consisted of validation team members, local project personnel and state Title III personnel.

The first question asked the participant to rank ten items on the following scale:

4 - excellent

3 - good

2 - fair

1 - poor

The averages across all ten training sessions are as follows:

Rank	Categories	Av. Rating
1	Consultant's Knowledge of Topic	3.6
2	Appropriateness	3.2
3	Handouts Distributed	3.0
4	Materials Presented	2.9
5	Information Presented	2.8
	General Evaluation	2.8
6	Length of Presentation	2.6
	Quality of Visuals	2.6
7	Method of Presentation	2.5
8	Activities Experienced	2.3

Fig. 1

2. See Appendix B.

Items 1 through 5 in Fig. 1, dealing with the appropriateness of the workshop average out to 3.05. It seems reasonable to us to portray the figure in this light since the remaining categories, 6 through 8, represent method of presentation which were constrained by the very short time line for preparation. (These items average out to 2.5)

A random comparison of regional responses indicates a high level of uniformity of response from workshop to workshop. This uniformity augers well for the standardized implementation of the validation process across the country, and as such fulfills a major objective of the training contract with SMA/ESD.

Also important are the findings relative to the participant's self-analysis in terms of how well he understood the validation concept prior to workshop participation, and then the increment of growth as a result of participation. See Fig. 2. The figures indicate that 62% of the participants categorized their knowledge prior to the training session as "poor" to "fair", whereas 75% indicated their understanding after the workshop was "moderately" to "greatly" improved. In light of the excellent credentials, the professional training and experience the participants brought to the training session, we find these figures interesting. One possible interpretation might be the newness of the content and the processes being undertaken. Another might be the distinctions made between validation as the process of verifying relevant documentation. In either event, the number of participants expressing feelings of growth and the responses to "appropriateness", "information presented", and "consultant's knowledge of topic", indicate that the validation concepts were heard and responded to affirmatively.

The actual figures are as follows:

"Prior to the training session my knowledge of the Title III validation process was:"

Responses	Number Responding	% of Total
Poor	75	29.6
Fair	83	32.7
Adequate	42	16.5
Good	46	18.1
Very Good	8	3.1
		100%

(N=254)

"As a result of attending the training session, I believe my knowledge, skills and abilities for conducting on-site validation procedures are:"

Responses	Number Responding	% of Total
Not Improved	6	2.4
Slightly Improved	57	22.4
Moderately Improved	113	44.5
Greatly Improved	78	30.7
		100%

(N=254)

Fig. 2

Readers will note the unusually high correlation between the pre-session "fair" and "poor" categories and the post-session "moderately improved" and "greatly improved" categories. Similarly, the pre-session 3.1% specifying "very good" knowledge does, by definition allow for much improvement, and the "not improved" (2.4%) figure bears out the correspondence. Please note attached charts (Page 9).

A concluding comment on the evaluation of this training may be in order to strengthen and facilitate the process for Year III. We would strongly recommend that consideration be given to designing the entire workshop program as an integrated experience, i.e., combine the regional meeting with the validation training. This would mean that personnel providing the standardization of the training nationwide need to be built into the entire instructional process rather than addressing only the Handbook/ Instrument on the second day (or the first day, depending where validation training can be most effectively and optimally conducted).

We also recommend that the entire workshop be evaluated. By obtaining participant responses to both days of training, a better total instructional package can be developed.

Fig. 4 portrays pre/post-test data prepared for 226 respondents. The bi-modal distribution in the pre-test is very likely explained by participants who were new to validation (and/or possibly by the fact that they did not have the validation information in sufficient time prior to the training session to have absorbed the concepts).

The pre/post-test questions were essentially informational as opposed to conceptual. The difference between averages, 53.03 and 79.36, pre and post respectively shows an overall growth of approximately 25%.

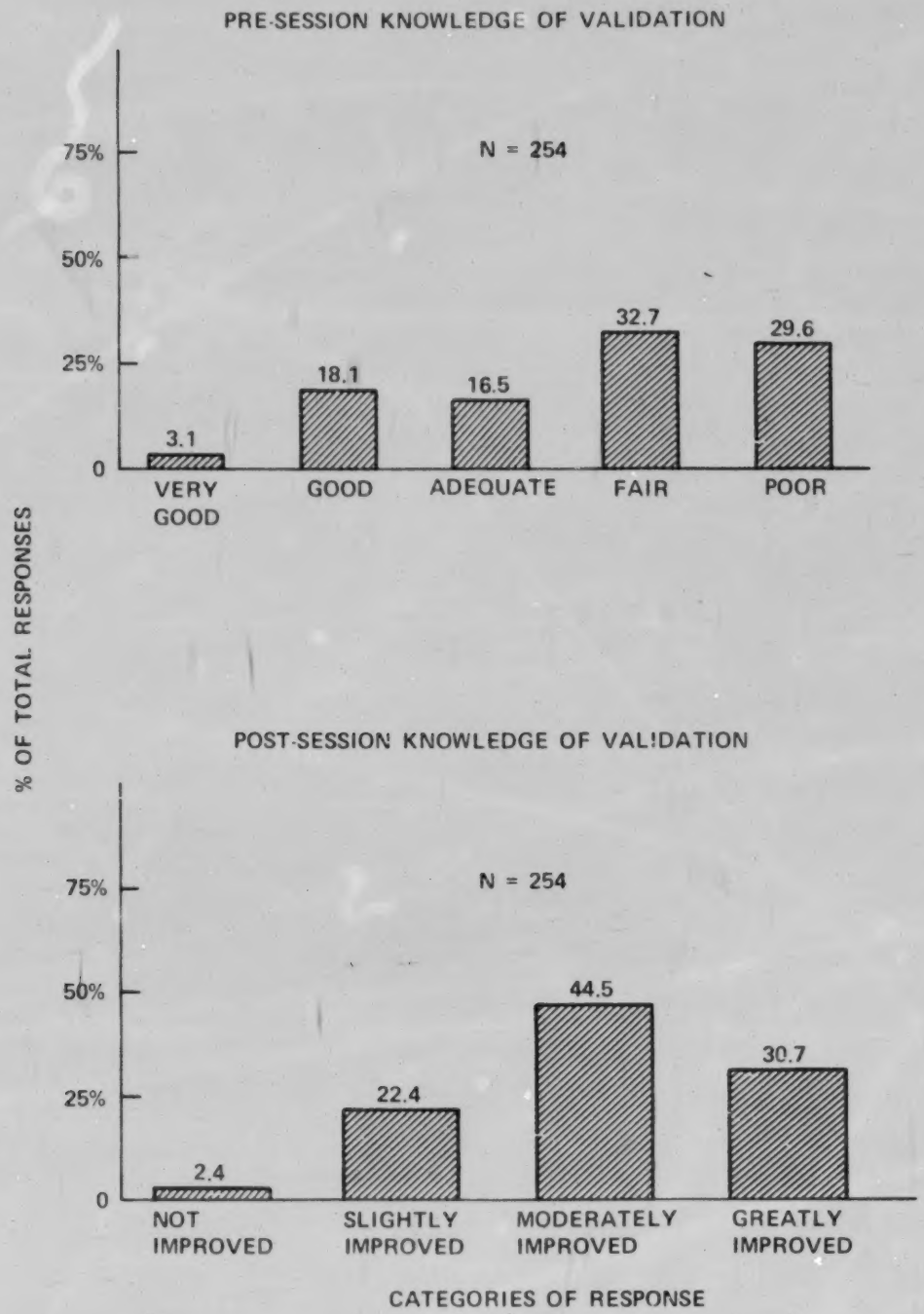


FIG. 3

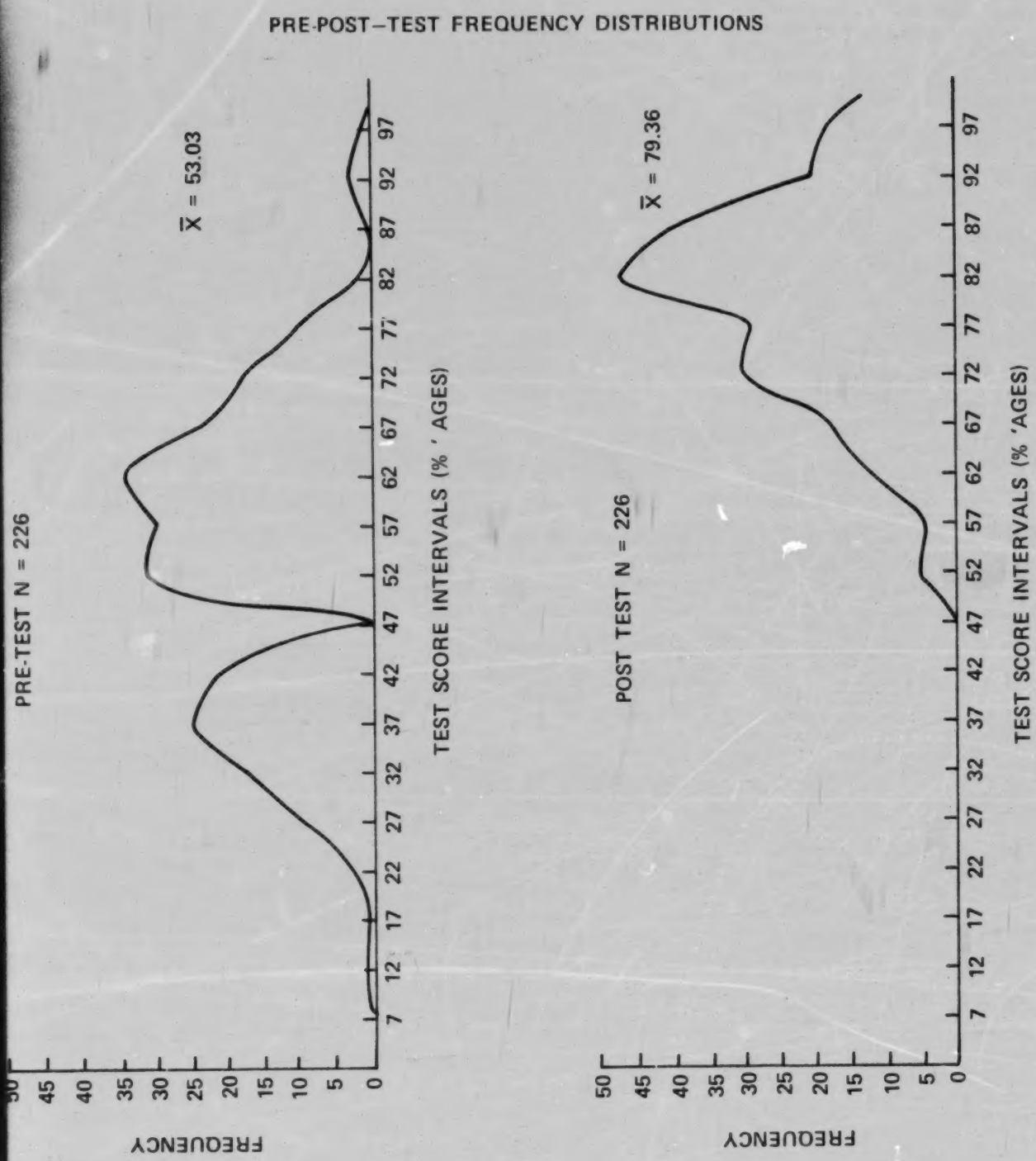

PRE-POST–TEST FREQUENCY DISTRIBUTIONS
PRE-TEST N = 226
$\bar{X}$ = 53.03
FREQUENCY
0
5
10
15
20
25
30
35
40
45
50
7 12 17 22 27 32 37 42 47 52 57 62 67 72 77 82 87 92 97
TEST SCORE INTERVALS (% 'AGES)
POST TEST N = 226
$\bar{X}$ = 79.36
FREQUENCY
0
5
10
15
20
25
30
35
40
45
50
7 12 17 22 27 32 37 42 47 52 57 62 67 72 77 82 87 92 97
TEST SCORE INTERVALS (% 'AGES)

FIG. 4

III LOGISTICAL PROBLEMS AND RECOMMENDATIONS:

There is no question in this writer's mind as well as in the collective mind of workshop participants (in this writer's judgement) that the 1974 Validation Instrument represents a substantial improvement and refinement over last year's. However, caution should be exercised before resting with this Instrument. There are still portions and/or specific questions in the Instrument which were still unclear or perceived as ambiguous in the training sessions and were subject to varied interpretations. Some of the problems will be addressed briefly here.*

There is the problem of "ownership" of the instrument. There were questions as with whom the locus of ownership resides. Therefore, it is recommended that:

1. identification of personnel should appear in the Instrument, and that
2. a review and revision panel be kept at a small maximum (12 at the most) with all of the expertises required to review, and revise the Instrument.

---

*The on-site visits are to be conducted in April and early May of 1974. Validators were requested to respond via the Validator Self-Analysis on the adequacy (or inadequacy) of the instrument in on-site application. A more fully-developed report will be submitted in full based on this important feedback.

There was the problem of lack of sufficient time to adequately field test the Instrument prior to training. Therefore, it is recommended that:

1. Year III revision begin at once such that a revised instrument can be produced in time for initial field testing and that
2. field test revisions be incorporated into the instrument before training

It became apparent as the workshops progressed that many of the validations proceedings were being assimilated by some state coordinators for inclusion in their own application-for-funding procedures. We, therefore, suggest that:

1. a panel be established to review these critical elements for validation which should be incorporated into state guidelines -- validation rests squarely on a standarized information gathering and assessment categorization base
2. that the procedures be standardize and made uniform throughout, thereby making the final results more immediately useful

With regard to some of the conceptual and theoretical motions underlying the Instrument (these will be addressed in breadth and depth in the final report), we would recommend that:

1. Uniform scaling be developed and applicable to all sections of the instrument. I.e., there was not, in effect, a scaled choice for Section II, Cost Information

that would allow for a value of something less than 3 and still allow the project to be validated.

2. all scaling terms, i.e., "persuasive", "conclusive", etc. should be applicable to all questions in all sections of the Instrument and that they be operationally defined.

3. some correspondence should be developed for validating other criterion sections by objective -- if objectives are the all-important criterion for judging the success (or lack thereof) of a project

Validation training clearly has to be improved. Some initial recommendations are as follows:

1. Engage prospective trainers early in the revision process

2. Allow sufficient time for a simulated training program to be developed: write objectives, conduct literature search, review and revise instrument based on Year I and II input, prepare initial training package, field-test training package in more than one site, etc. Six months should be ample for this.

We felt the data presented in the Guide was well received. We heartily recommend that this information be reviewed for inclusion in the Handbook's Year III revision both in terms of the mechanics of the on-site visit, and for Instrument revision and improvement. Clearly the state coordinators

need to be surveyed for their recommendations on ways to improve the logistical details since attention to this type of detail is as critical to effective validation as the validity of the Instrument itself!

We recommend that a mini-Handbook be prepared for State and local project personnel to be sister documents for the materials for validators. Both documents clearly required numerous interfaces and internal consistency, but unless the local project people have better instructions and a longer lead time in gathering supportive documentation ("making their case") the process will not work as efficiently as we believe it can.

We also recommend that a Validator Bank be established with the names approved for inclusion chosen on a discriminating basis including, among others, the following factors:

1. the state coordinator's analysis of the validator's efforts on site, and the written results of that effort
2. the analysis of the validator's written criteria (i.e., the section for which he/she was responsible)
3. feedback from project personnel
4. the assessment of the validator's self analysis on site, etc.

At this point it is appropriate to note that the credentials and experience of those individuals chosen to serve as validators were of an unusually high order. It is reasonable to assume, therefore, that is these gifted professionals cannot make a success of the effort this first time around, then in all likelihood no other group of professionals could either!

Finally, we believe that a procedure needs to be developed and adopted, as a result of the Year III revision of the Instrument, (and the on-site procedures!) wherein the Division of Supplementary Centers and Services personnel can have test data on how completely personnel selected to be trained as validators actually know and can demonstrate the required assessment skills! As with the developing procedures for the role of educational auditor there needs to be an agreed-upon competency base below which professional personnel not be invited to serve as validators.

Competency data could be generated from several sources:

1. testing (pre and post)
2. assessment of personnel on site
3. analysis of written validation reports

Part III

Recommendations

"Validation" represents fundamental theoretical and methodological departures from the ways in which educators have traditionally evaluated programs. It is our contention that validation deserves and requires a more comprehensive and internally consistent developmental plan if it is to be optimally responsive to the needs of the educational consumer. While giant strides have been made with this initial validation effort, it is clear from the feedback received nationally that more intensive effort needs to be applied to the testing and resultant validation of the Instrument, and for comprehensive planning for Year III and following.

Since the end product of the entire validation effort is the increased national assimilation of educational practices validated as exemplary and effective, it seems apparent that a strenuous effort needs to be directed to the continuing and competent revision and implementation of the total process. Only exhaustive and coordinated efforts at Instrument and Handbook revision, and carefully planned and executed training workshops will result in the quality product required by the public for accountability, the Congress for continued funding, and the integrity of the process required for assimilation by the nation's educators.

The following recommendations are made relative to IVD planning, Instrument revision, and training procedures based upon SMA's extensive involvement in the total validation process. These recommendations reflect our critique of the Validator Self-Analysis Forms, the validator-prepared Reports, the team's participation in the Instrument's preparation, and the evaluation of the conduct of the training workshops.

Readers are respectfully directed to Section I of this Report for detailed statements of issues and implications for Instrument revision, as well as Section III-B, and to the Appendices for a proposed time schedule for the development of Year III validation procedures.

A. Validation Planning and Training Procedures

1. We strongly recommend that a representative Panel be selected to work with the training and Instrument revision contractor: that the Revision Panel contain twelve members stipulated by name, and that the Panel not meet to take action on the validation process unless a quorum is in attendance; and, that, further, the Panel "own" the results of the revised procedure including both the Year III Instrument, and the operational procedures themselves.

2. That the Review Panel's advisory duties would include, among others, the following tasks:

    1) arbiters of a Section's point value, and of the value of each question with a Section

    2) assisting in identifying appropriate sites for field testing the revised Instrument

    3) as an advisory group to respond to the contractor's recommendations, and/or their own, for the inclusion of new questions/sections

    4) the approval of the revised Handbook, training procedures, and the Instrument

3. That USOE/DSCS/OSLEP personnel hire one firm as prime contractor for Year III revision activities, and that said firm be charged with the responsibility of planning and implementing Year III activities including the revision of the Validation Instrument

4. That presentations at the ten (+-) workshops be conducted by one contractor for the sake of uniformity, and, that the training or "content" portions of the workshop be interfaced with state/validator organizational concerns. This proposed balance in presentations will allow discussion immediately of both the on-site logistical details, and the demands theoretically and Instrument-wise of the validation undertaking.

5. That a simulated learning package be prepared -- and time to do so -- for use in establishing coder reliability for practice and testing in the training workshop, and, further, that as a critical aspect of the field-testing of the Year III IVD Instrument and "back-up" teams review the findings of the earlier team, and that both sets of results and scores be matched for the identification of coder discrepancies. Such follow-up team validation assessments would need to be conducted on a shorter term and randomized basis.

6. The entire workshop should be evaluated, and not just the performance of the training contractor. We further recommend that the state coordinators elicit feedback on the entire selection and on-site visitation procedure. We also strongly recommend some form of pre/post testing (of those selected to be validators) for feedback on what their understanding of the validation task may be, and for the identification of problem areas to be addressed in the workshops as a result of the pretesting.

7. That state coordinators be encouraged to conduct a post mortem session with both validators and project personnel for feedback, and that such capta be related to IVD personnel in Washington

8. Project personnel with practices to be validated should be invited to attend the Validator Training Workshops

9. That the revision of the Handbook be directed to:

    1) the preparation of a complete glossary of all IVD terms in the Instrument and that said terms are operationally defined

    2) the inclusion of a section on the theory of validation as the proposed rules for evidence review, and the critical distinctions the IVD process makes between evaluation and validation

    3) revisions in the procedures for team member selection and assignment

    4) more completed and sensitive directions for local project personnel in the completion of the blue sheets

    5) the functions of the team leader, and complete details on the recommended team interaction process

    6) the advantages and disadvantages of team observation of practices

    7) a discussion of possible validator biases

8) a comparison of the roles of IVD validator and educational auditor

9) detailed instructions on preparing the Validation Report

10) the procedures for conducting the on-site visit including a proposed time schedule

11) the specification of the state coordinator's responsibilities

12) the specification of the regional coordinator's responsibilities

10. That pertinent technologies and methodologies to increase dissemination/exportability potential be included, and with specific reference to:

1) instrumentation addressed to decision-settings, evaluation formula, and planning models

2) cost formulations on a per pupil per instructional hour basis

3) cost conversion scales for the geographic comparison of costs for potential consumer districts

4) instrumentation addressed to identifying the qualities required for effective leadership in project replicability, i.e., how is the charismatic leader's behavior to be analyzed in terms of actions essential to success in the consumer district

5) a system to code "practices" against consumer needs, i.e., by academic area, size, staff/student ratios, costs per pupil, futures orientation, demographic descriptions, etc., all directed to providing strategies for educational change

11. That adequate time be allowed for both the field testing of the revised Instrument, and for the training of local project personnel in the preparation of documentation (for use on the blue sheets)

12. That a "Mini-Handbook" be prepared for state and local project personnel citing case studies and other illustrative data of what constitutes acceptable documentation, the proper form for objectives, appropriate examples of evaluation designs, management instruments, and testing procedures, etc. Such a Mini-Handbook would greatly facilitate the completion of the "blue sheets", and would facilitate the on-site review of documentation

13. Validator team assignments should not be determined until after the training workshop is completed -- particularly with respect to the selection of the chairperson

14. That a Validator Bank be established in order that trained and experienced personnel will be on record for use by the states as the IVD process grows and, further, that a procedure be developed for certifying said validators employing competency data from at least the following sources:

    a) pre and post testing

    b) assessment by state personnel of their work on site

    c) analysis of their written validation reports

15. Procedures for identifying potential validators should be uniform throughout the ten regions

16. That the SMA proposed time schedule be adopted with such modification as may be necessary

17. That instructions for formating validated project findings for submission to ERIC be included in the revised Handbook

B. Instrument Revision

1. That the Instrument be thoroughly revised as a result of USOE/DSCS, SMA, validators, PNAC and NASACC criticisms

2. That the blue sheets be correspondingly revised

3. That the Instrument be revised to show "profile" data that can be optically scanned and computer tabulated for rapid classification

4. That the Instrument request data indicating both the need, and the state-wide priority ranking of the need to which the project is an effective response

5. That the detailed criticisms synopsized by the SMA analysis of the Validator Self-Analysis Form be addressed in detail. Please see Section I of the Report.

APPENDIX A

January 16, 1974

Prospectus

Validation Handbook Revision;
Validation of Instrument Revision;

Preparation of Content for and Conduct of Year II Training - Implementation Workshop and On-Site Visitations

Submitted to:

Mr. Thomas Burns, Deputy Associate Commissioner
U.S. Office of Education
Bureau of Elementary and Secondary Education
Washington, D.C.

Submitted by:

The Educational Systems Division
SCIENTIFIC MANAGEMENT ASSOCIATES, INC.
Route 130 and Nicholson Road
Gloucester, New Jersey 08030

Introduction and Rationale

"Validation" represents fundamental theoretical and methodological departures from the ways in which educators have traditionally evaluated programs. It is our contention that validation deserves and requires a more comprehensive and internally consistent developmented plan if it is to be optimally responsive to the needs of the educational consumer. While giant strides have been made with this initial validation effort it is clear from the feedback received nationally that more intensive effort needs to be applied to the testing and resultant validation of the Instrument, and for comprehensive and technologically competent planning for Year II and following.

Since the end product of the entire validation effort is the increased national assimilation of educational practices validated as exemplary and effective it seems apparent that a strenuous effort needs to be directed to the complete and competent revision and implementation of the total process. Only exhaustive and coordinated efforts at Instrument and Handbook revision, and carefully planned and executed training workshops will result in the quality product required by the public for accountability, the Congress for continued funding, and the integrity of the process required for assimilation by the nation' educators.

SMA is in a uniquely advantageous position to undertake the proposed tasks outlined in this Prospectus. First, SMA/ESD personnel have been deeply involved in the generation of the first Instrument. Second, SMA/ESD personnel prepared and conducted the national workshops for Instrument utilization. An essential aspect of the conduct of the workshops was the collection and analysis of detailed responses to both the mechanical procedures in the validation on-site process, and to needed changes in the

Instrument. Critically, SMA/ESD personnel are the only individuals who have been involved in the entire process from procedures development to final evaluation for Year I. The participant critique of SMA/ESD personnel in the Workshops indicates a high confidence level for continued participation in the developmental effort. Additionally, the corporation's professional assets and capabilities make SMA a logical choice as prime contractor for the expanded development and improvement of the national validation effort.

Cognizant of the need to maximize producer/consumer interaction through the vehicle of educational validation, SMA proposes to deliver the following products and services.

The following sections reflect SMA's professional judgements on the need for Handbook and Instrument revision and for needed improvements in implementation. These six sets of items represent our responses to extensive feedback from local, state, regional and federal personnel during the conduct of fourteen (14) training sessions coast to coast. The responses are also reflective of those needed logistical details that would only be apparent to a contractor faced with training personnel in the use of the Handbook.

We submit these items for consideration fully aware of the cost and the thousands of professional man-days of effort committed to the process to date. Our criticisms of existing processes are for the sole purpose of improving the validation process and for expediting the assimilation of cost-effective educational practices.

I. Handbook Revisions: Scope of Services and Products for Delivery

A. Supervising Handbook Revision Task Force

SMA will assist in the construction of a representative panel of twelve or more people to oversee the revision of the Handbook. The Task Force working in close relationship with SMA would serve as:

1. arbiter of a section's and a question's relative value using an appropriate weighting procedure (Q Sort, Delphic survey, etc.)
2. an arm in selecting appropriate sites for field tests of the phase II revisions
3. assistance in identifying teams to conduct document analysis as a validation of previous team findings
4. the approval body for the successive revisions of the Handbook and Instruments

Acting as the executive arm for federal personnel, SMA will make suggestions relative to the composition of the Task Force, specify Task Force duties, and oversee the scheduling and conduct of meetings.

B. Logistical/Mechanical Revisions

SMA personnel will revise the Handbook's procedures to reflect the actual implementation of Year I and II validation proceedings, as well as to include suggestions for improving implementation for Year II and III. It is apparent to us that the improvement of the

mechanical details of the total validation effort is just as critical as revised and improved instruments. In the rewriting of the Handbook's white pages for Year III, careful attention will be directed to:

1. the preparation of a complete Glossary
2. the inclusion of a section on the theory of validation as the rules for evidential review, and the critical distinctions between validation and evaluation
3. revisions in the procedures for team member selection and assignment
4. more complete and sensitive directions to local project personnel in the completion of the Self Analysis Form (the blue sheets)
5. specific and scheduled training workshops for those serving as validators
6. the functions of the team leader, and a complete itemization of the team interaction process
7. the advantages and disadvantages of team observation of practices
8. potential validator biases
9. the role of validator compared with the role of educational auditor
10. instructions for preparing the validation report
11. the procedures for on-site validation including a time table
12. a more coherent outline for the final Validation Report, including instructions to local and state personnel on required content and format
13. the responsibilities of the state coordinators
14. the responsibilities of the regional coordinators, and,
15. detailed attention to the content and scheduling of the training workshops

C. Instrument Revision

SMA personnel will present to the Validation Handbook Revision Task Force a revised Validation Form for on-site use based on Year I and II IVD training. This revision will reflect the multiple changes needed to give the Instrument greater strength, academic credibility, and, after field testing, validity and reliability. This revision, once approved by the Task Force, will be employed by multiple teams using the same documentation to check for coder reliability. Revisions resulting from this internal validation will then be proposed for inclusion as the revised Instrument for Phase III.

D. Project Nomination Form

As a concurrent activity, SMA personnel will revise the blue sheets to correct identified weaknesses in Phase I, and to be compatible with the revision of the On-Site Validation Form. Revisions will include formating, clarification of the questions, and examples of desired materials.

II. Preparation of Content for and Operation of Regional Training Workshops

A. SMA will schedule workshops for state and project personnel, those to serve as validators, and BESE personnel.

Content for these workshops will reflect all revisions and new procedures. Workshops for all levels of personnel can be held in the same time period. This overlapping of levels of personnel will result in improved documentation and accelerated on-site visits. Additionally, this multiple targeting approach

to conducting the workshops will accelerate the dissemination of effective practices regionally, and the inclusion of validation procedures in state application forms for funding new projects.

B. SMA will prepare in narrative form, for inclusion in the Handbook, the desired format for the conduct of the workshops interfaced with the responsibilities of federal, regional and state coordinators.

III. Special Work Tasks

SMA personnel will propose for the Task Force's consideration pertinent technologies and methodologies to enlarge the dissemination and importability potential of effective educational practices. These additions would include:

a. instrumentation addressed to decision-settings, evaluation formula, and planning models;
b. cost formulations on a per pupil per instructional hour basis;
c. cost conversion scales for geographic comparison;
d. a package of simulated training experiences for validators to be used as a screening device in identifying potentially low-effective validators;
e. the procedures to establish a Validator Bank of certified professional personnel including criteria for selection;
f. instrumentation addressed to identifying the qualities required for effective leadership in project replicability;
g. a system to code "practices" against consumer "needs", and,

h. a detailed cost analysis comparison of how to train the maximum number of personnel on all levels at minimum cost

IV. Assessment and Classification of Year I Validated Projects (Practices)

SMA will:

a. identify response problems
b. classify reports by category and type
c. coordinate findings for computer access with an appropriate computer installation (e.g., Kentucky Title III Project)
d. make recommendations on dissemination format, and
e. prepare ERIC Abstracts

V. The Preparation of Validation Guidelines for State Application Procedures

SMA proposes that key elements of the documentation process be modified for inclusion in State application procedures. SMA will prepare such guidelines for adoption by interested states. The adoption of these guidelines will facilitate the validation of educational practices and will expedite the matching of effective practices to particular learning needs.

VI. Publications Production

SMA's publishing subsidiary Scientific Management Publishers, is capable of producing all printed matter required, at competitive prices, and on a very short time schedule. Possible options for

publication:

a. the revised Handbook

b. supplementary materials for the conduct of the workshops

c. materials for the Educational Fair

d. materials to mail to school districts requesting additional project data

VIII. Time Line and Costs

SMA is proposing the adoption of the services specified in Sections I through V as the basic contract package. It occurs to us that these are the essential components of the total validation process, and that all of these elements need to be addressed as a synergistic whole. Section VI provides for publication services. Costs would be determined by the size and format of the material in question.

Costs for services are determined by computing man days of effort, overhead, general and administrative expenses, materials and supplies, and profit. Specific costs will be prepared for submission along with a complete proposal responsive to those items negotiated for inclusion.

The formal proposal will include a PERT network, including subsystems for materials classifications, data processing for computer access, and printing.

Proposed SMA/ESD Delivery Schedule

MONTH	SERVICE(S)	DUE DATE	PRODUCT(S)	DUE DATE
March '74	Construct Validation Panel Schedule Panel Meetings	1 7	Submit list of Panel's duties; Submit procedures for weighting questions and Sections	15 30
April	Revise Handbook (white & blue pages Categorize/Classify Phase I Reports	1-30	Submit Phase I Report on Classification of Practices	15
May	Revise Handbook (yellow & green pages)	15	Submit revised Handbook to Panel	15
June	Field Test Phase II revision Set up Computer Access procedures	1 30	Submit corrected copy re field tests	30
July	Validate Phase II revision Identify Workshop participants	1 15	Submit corrected copy re validation	30
August	Prepare content for Phase II Workshops; Critique and revise	15	Submit Computer Access Procedures	30
September	Schedule workshops Prepare format for workshops	1 15	Submit Computer program Submit workshop schedule	15 15
October			Submit workshop format	30
November	Conduct workshops	1-30		
December	Conduct workshops	1-15	Submit report on conduct of workshops	30
January '75	(Conduct of Validations)	1-30		
February	Collect, analyze, classify validation reports	1-20	Submit classifications	21
March	Prepare abstracts	1-30	Submit materials for Ed Fair; Submit Abstracts	30 30
			Submit materials for publication	30

APPENDIX B

TABLE OF CONTENTS

	Page
Pre-Test	1.
Training Schedule	5.
Team Procedures	6.
Team Interaction and Preparation of Final Validation Report	10.
On-Site Validation Checklist	11.
Proposed Daily Schedule for On-Site Visit	12.
Problem on Cost Information	13.
Procedures for Validation of Evidence on Effectiveness/Success	14.
Problem on Effectiveness/Success	16.
Summary Rating Worksheet	21.
"Evaluation" (4,a-i) TABLE OF VALUES	22.
General Evaluation Design	23.
Evaluation Design Checklist	24.
Basic Management Design	27.
Control Group Assumptions	35.
Factors on Internal and External Validity	36.
Interview Techniques	38.
Indices of Specific Components of Title III Projects	44.
Unobtrusive Measures	46.
Needs Assessment Scheme	48.
Dissemination and Utilization Strategies	49.
Catalog of Statistical Tests	51.
Model Project Final Report	61.
Validator Self-Analysis Form	69.
State Codes	70.
Bibliography	71.

PRE-TEST ON THE HANDBOOK FOR VALIDATION OF EDUCATIONAL PRACTICES

1. It is the team chairman's responsibility to
   (a) Prepare the narrative report for each criterion section (a)_____
   (b) Coordinate preparation of preliminary reports but write no report (b)_____
   (c) Serve as an expert in one of the criterion sections and write only his section's report (c)_____
   (d) None of the above (d)_____

2. The review and rating of the Applications for Validation are conducted by
   (a) The SEA (a)_____
   (b) The USOE (b)_____
   (c) The ESEA Title III Advisory Board (c)_____
   (d) Jointly by SEA and Title III Advisory Board (d)_____
   (e) None of the above (e)_____

3. Validators serving on a validation team are chosen by
   (a) The Regional coordinator from their Region (a)_____
   (b) The USOE/Title III Office (b)_____
   (c) Their own State Coordinator (c)_____
   (d) The validation team chairman (d)_____
   (e) None of the above (e)_____

4. The minimum number of validators serving on a validation team is
   (a) 2 (a)_____
   (b) 3 (b)_____
   (c) 4 (c)_____
   (d) 5 (d)_____
   (e) Dictated by the size of the project (e)_____

5. The validation team, once selected will
   (a) Validate only one project in their own state (a)_____
   (b) Validate only one project in a state and region other than their own (b)_____
   (c) Validate only one project in a state other than their own but in their region (c)_____
   (d) Validate more than one project in a state other than their own. (d)_____
   (e) None of the above (e)_____

6. The responsibility for developing a resource pool of Validator/Specialists from which to select a validation team rests with
   (a) USOE (a)_____
   (b) Regional Title III Coordinator (b)_____
   (c) CSSO (Chief State School Officer) (c)_____
   (d) The State Title III Coordinator (d)_____
   (e) None of the above (e)_____

7. If a nominated project has a dual focus, e.g., an early childhood project for handicapped children, the team membership should
   (a) Be kept at a minimum of three (a)_____
   (b) Be increased to four, or more, if needed (b)_____
   (c) Be kept at three but have one member with expertises in more than one area (c)_____
   (d) Add additional members as needed so the chairman does not have to act as a validator (d)_____
   (e) None of the above (e)_____

8. The SEA's Title III staff members may serve on a validation team if he (she)
   (a) Is a part-time employee of Title III (a)_____
   (b) Validates in a state other than his (her) own (b)_____
   (c) Validates in a region other than his (her) own (c)_____
   (d) Serves only as an observer (d)_____

9. In order to be eligible for validation team membership a member must
   (a) Have participated in last year's validation training (a)_____
   (b) Participate in this year's training (b)_____
   (c) Have been nominated by his regional coordinator (c)_____
   (d) All of the above (d)_____

10. Each team member, after leaving the project site will
   (a) Be sent the final report from the team chairman for sign-off (a)_____
   (b) Send his own section narrative report to the Title III coordinator of his (her) state (b)_____
   (c) Send his own section narrative report to the state coordinator in which the project is located (c)_____
   (d) Send his section report to the team chairman for inclusion in the final report (d)_____
   (e) None ofthe above (e)_____

11. Validation for any given project is accomplished only by objectives
   (a) In all criterion sections (a)_____
   (b) In all but one criterion section (b)_____
   (c) In only one of the criterion sections (c)_____
   (d) None of the above (d)_____

12. Any well-stated behavioral objective should include
   (a) The name of the community (a)_____
   (b) The total number of the student body (b)_____
   (c) The performance level (c)_____
   (d) All of the above (d)_____

13. The final validation report is to be submitted to the USOE
   (a) Within 10 days of the on-site visit (a)_____
   (b) Within 2 weeks of the on-site visit (b)_____
   (c) Within 3 weeks of the on-site visit (c)_____
   (d) Within 1 week of the on-site visit (d)_____
   (e) None of the above (e)_____

14. Once it has been agreed to by the validation team that a project meets the minimum number of points for all criteria the project is
    (a) Accepted for national validation (a)_____
    (b) Accepted for state validation (b)_____
    (c) Accepted for USOE validation (c)_____
    (d) Accepted to participate in the National Ed Fair (d)_____
    (e) None of the above (e)_____

15. It is the local project staff's responsibility to
    (a) Provide prject's financial records (a)_____
    (b) Provide office materials needed by the team (b)_____
    (c) Have available all dissemination materials (c)_____
    (d) Prepare assurance by the local superintendent regarding project continuation (d)_____
    (e) All of the above (e)_____

16. The SEA or LEA may submit project summary data for approval directly to the
    (a) USOE Dissemination Review Panel (a)_____
    (b) USOE Division of Supplementary Centers and Services (b)_____
    (c) Regional Title III Coordinators (c)_____
    (d) None of the above (d)_____

17. The principal source of the items in the Handbook is
    (a) NACSCS (National Advisory Council on Supplementary Services and Centers) (a)_____
    (b) USOE (b)_____
    (c) NASACC (National Ass'n of State Advisory Council Chairmen) (c)_____
    (d) SEA's (d)_____
    (e) All of the above (e)_____

18. "A validated practice that is feasible to communicate to other schools---with similar needs and environments" is said to be
    (a) Validated (a)_____
    (b) Innovative (b)_____
    (c) Reliable (c)_____
    (d) Exportable (d)_____
    (e) None of the above (e)_____

19. In the event a validator,or the team as a whole, felt that there was insufficient documentation to validate the project, the team (or validator)
    (a) Should ask if there are more data (a)_____
    (b) Should immediately invalidate the practice (b)_____
    (c) Should causus to determine if further validation is warranted (c)_____
    (d) None of the above (d)_____

20. The statement "as to the innovativeness of the project..." (Section D of the Final Report) must be based on
    (a) 2/3 agreement of the team (a)_____
    (b) This practice occurring in not more than 5% of the state's school districts (b)_____
    (c) The best professional judgments of the team as being uncommon, creative and original (c)_____
    (d) All of the above (d)_____

TRAINING SCHEDULE AGENDA

Day One

8:30 - 9:30	USOE Orientation
9:30 - 9:45	Introductory Remarks on Validation
9:45 - 10:15	Slide/tape on "The Case for Development"
10:15 - 10:30	Break
10:30 - 11:00	Pre-test and scoring
11:00 - 12:00	Handbook/Guidebook Orientation
12:00 - 1:00	Lunch
1:00 - 1:45	Special Compututional Problem in Handbook Regarding Cost Information
1:45 - 2:30	Effectiveness/Success Amplification
2:30 - 2:45	Break
2:45 - 3:45	(1) Problem Session for Validators (2) Special Session with State Coordinators and Team Chairman
3:45 - 4:30	Discussion

Day Two

8:30 - 12:00	Final Validation Report Writing

I. Team Procedures

Introduction:

The special task confronting the validation team is that of assessing the utility of the information presented in terms of its credibility as evidence in validating an exemplary practice. The burden of the task, therefore, is to weigh the multiple evidence presented in terms of the judgments required by each of the questions. Clearly this is not the traditional approach to evaluation generally practiced in the nation's schools. Rather, the purpose of the validation effort is to verify the credibility of the project's practices and the reports on those practices.

The very fact that a practice has been nominated constitutes its innovativeness and success on the state level. It is, therefore, the team's responsibility to determine if what the project personnel said was happening, was in fact happening, on the basis of the tangible evidence submitted for review.

This evidential approach is critical to the practice's adoption since it is on the basis of the team's review and summary of the written documentation that other districts will have access to the information.

It is clearly necessary, therefore, that the team review the evidence both individually and collectively. The team chairman has been instructed to allow time for total team consultation prior to the preparation of the Validation Report in order that there may be a good interchange of data on the total operation of the practice within its institutional setting.

II. Instrumentation:

The On-Site Validation Report Form has been purposely designed as a self-contained unit. No additional interpretive nor recording data is needed or desired. Each question is self-contained, requires its own data base, and is separately rated. Clearly the individual validator will be responsible for setting his own parameters for rating judgments based upon the availability of material. Because this is so, and the subjective judgment of the validator is the only basis for making response, it is necessary to document the basis on which the decision was made, and to cite critical evidence. It is also clearly evident that the validator must make the judgment as to the adequacy of the data presented for review.

As indicated in the SMA Guidebook for On-Site Validation Procedures, it is obviously essential that each validator be able to describe in his own words the major objectives for the practices being validated. The major objectives are the sole basis of the

questions, and all judgments must be made with the validation of the objective as the single concern. Where a validator feels strongly that notation should be made of particular processes utilized in the success of the practice, such information should be so noted in the narrative comment under Validation of Evidence. Recommendations for weighting responses:

For each question we propose four (4) headings for reflection by each individual validator, and then for the team as a whole:

I. Within the parameters of the question are the exhibits adequate to cover the relevant variables? (Please Note: As a framework of reference, variables may be classified under three headings - 1) behavioral, 2) instructional, and 3) institutional. The behavioral categories include the three domains - cognitive, affective and psychomotor; the instructional variables include organization, content, method, facilities and costs; and the institutional variables include profiles on students, teachers, administrators, educational specialists, families and community.)

II. Does the documentation (exhibits, evidence, etc.) support the nomination of the practice for national visibility? Is the documentation inclusive enough to allow for adoption?

III. Is there high coder reliability? Has each validator requested and received feedback from his team mates on their weighing of the questions in his section? Is there reliability of response?

IV. Does the team's summarization of the practice in question provide all the critical data needed for the adoption of the practice?

The team chairman might find it useful to adopt a procedure such as the matrix below for weighing the team's response to each question, and for making judgments as to the adequacy of the supportive documentation and exhibits.

Validation Check Sheet

Section # _____

Question # _____

a paraphase of the question:

Indicators

1. Materials cover relevant variables?
2. documentation adequate?
3. high coder reliability?
4. validation Report inclusive of all critical data?

Other considerations:

In order to regulate to the maximum degree possible the conduct of the on-site validation procedures the following suggestions for data analysis are being put in each validator's hands. As earlier noted, however, the "standards" within which the validator determines a number weighting for the question under scrutiny is totally dependent upon the scope and extent of the materials available on the day of the on-site visit.

A. EFFECTIVENESS/SUCCESS

The critical concern for this section is the extent to which the practice's objectives have been achieved and/or the learner's performances improved. Validators will want to consider:

1. extent objectives are critical to understanding success of the practice
2. extent performance levels are challenging and realistic
3. congruency between objectives and related activities
4. appropriateness of test selection
5. reliability of test administration
6. range and variety of instruments employed
7. appropriate data treatment procedures - descriptive, analytical, inferential, comparative
8. relevance and imagination of evaluation design

B. COST INFORMATION

This section posits as the standards against which effectiveness judgments are to be made the parameters of the data provided by project personnel. In short, these judgments can only be made in response to the performance levels achieved compared with the costs per pupil. It is possible that the validator might be assisted in this task by requesting that project personnel prepare cost figures on a per pupil per instructional hour basis. If this is not possible then the validator must estimate the performance levels achieved against the expended costs.

The validator will also want to carefully review the costs presented, and to make some judgments as to the completeness of the data. Where irregularities occur the validator may request primary sources.

C. EXPORTABILITY

The following considerations might prove helpful in responding to this section:

1. Will the practice be continued? Is the evidence for continuation encouraging?

2. Is there a high relationship between the local school district's use of the practices and the needs of the state at large?

3. If applicable, is there evidence of support by key constituencies?

4. Is planning, management and dissemination information adequate, clear and replicable?

5. Are critical processes and procedures well documented and critiqued?

6. How adequate was the identification of problems and the procedures for their resolution?

7. Will the data submitted by project staff, supportive and/or critical S.E.A. documentation, and the Validation Report of the team, along with attachments, serve the critical information needs of adopting districts?

Team Interaction and The Preparation of The Validation Report

The team is responsible for completing all the questions in the On-Site Validation Form in the Handbook. It must be emphasized that each question requires both an explanation for the number weight assigned, and the citationoof evidence reviewed in making the decision. The total team's responses will be included in the Final Report to be sent to Washington. The name of the validator responsible for each section of the Report is to be clearly identified.

Section C of the Report is the narrative summary for each of the three sections on Effectiveness/Success, Cost and Exportability. This narrative description of the practice's objectives, operation and evaluation must be comprehensive enough to provide an interested school district with all the information needed for adoption.

In the process of preparing both Section C and the On-Site Validation Form the team will need to:

(a) complete each individual section and write a narrative summary of findings.
(b) meet as a total team to weigh each of the other two sections of the Handbook.
(c) review as a total team coder reliability for each section, and then make response-weight comparisons with the team member assigned to that particular section.
(d) discuss areas of coder discrepancy.
(e) seek, where desirable, additional data on the issue under contention.
(f) agree to a point total for each section in the order prescribed in the Handbook.
(g) team members disagreeing with the majority opinion may prepare a dissenting report for inclusion in Section C.
(h) the team chairman collects the individual narrative summaries on each section, prepares such introductory data as required, and drafts the final written report.
(i) the rough summary is shared with team members and each team member signs-off on the rough draft.
(j) the rough draft is shared with project personnel as the essential content of the exit interview.
(k) the chairman takes the team-approved rough draft and prepares a final typewritten report.
(l) the chairman mails a copy of the Final Report, along with all three On-Site Validation Form responses, to the State Coordinator within 2 weeks of the visitation.
(m) each team member fills out the Validator Self Analysis Form at the conclusion of the on-site visit and mails the form to U.S.O.E.

ON-SITE VALIDATOR CHECKLIST

Note: It is recommended that this Checklist be utilized on the evening prior to the all-day validation effort. Any negative responses should be addressed before the validation activity gets underway.

	YES	NO
1. Have you read the Validation Handbook entirely?		
2. Have you read, and do you have in your possession the the following documents?		
a. the proposal abstract		
b. the original proposal(s)		
c. copies of project evaluation instruments and results		
d. the complete Application for Validation		
3. Are you familiar with:		
a. the data analysis procedures used?		
b. names and qualifications of personnel involved in test administration?		
c. educational/instructional materials produced by the project?		
d. the State's application and project evaluation procedures?		
e. the State's on-site financial and audit reports on the project?		
f. learner and community characteristics?		
g. information on the school system?		
h. critical educational needs specified by the project?		
i. the relationship between project activities and project objectives?		
j. the evidence noted by the project related to the three specific criteria: effectiveness/ success, cost and exportability?		

PROPOSED DAILY SCHEDULE OF ON-SITE VISIT

Evening of arrival

PM		Team introduction (The team, state and local project personnel) secure and review team member folders
		Review team member and total team responsibilities
		Complete On-Site Validator Checklist (and address any areas of need identified)

Day on-site validation

AM	8:00	Meet all project personnel, and secure work space
	8:10	Analyze quantity of data to be reviewed (printed materials, visuals, etc.)
	8:20	On the basis of the data to be analyzed schedule the remaining work day (observations, interviews and review of written materials)
	8:30	Observation of the educational practice(s) (if appropriate)
	9:15	Review of project data utilizing The Handbook's On-Site Validation Form (yellow pages)

Lunch

PM	1:00	Continued work on On-Site Validation Form
	3:00/3:30	Team meets to reach agreement on each section (criterion)
	3:30/4:00	Each team member prepares a rough draft narrative summary of his section
	4:00/4:30	(1) Team meets with project/state personnel for exit interviews (2) Team prepares conclusions and recommendations including minority reports (Section C) (3) Team completes a statement as to the innovativeness of the project as viewed by the team (Section D) (4) Team chairman submits rough draft of final report to team members for sign-off (Section E-3)
	4:45	Team members fill out Validator Self Analysis Form and place in mail.

PROBLEM: With reference to question number 3, page 85.
During the 1970-71 school year, 3,181 students and 121 teachers participated in project field activities. Each student spent an average of 13.7 hours in these field activities. These students were from grades 1, 3, 5, 7 and 9. Their classroom teachers each spent 12 hours in in-service activities.

Given this information, compute the average total number of hours per learner served.

Computational Process:

A. Find the total number of learner hours (round to nearest whole #)

   Students:

   Teachers:

   Total:

B. Find the total number of learners

   Students:

   Teachers:

   Total:

C. Find the average total number of hours per learner (round to nearest tenth)

   Total number of learner hours/total number of learners

D. ANSWER: 13.6

PROCEDURES FOR VALIDATION OF EVIDENCE

ON EFFECTIVENESS/SUCCESS

(Pages 65 through 81 in Handbook)

Step 1: Reproduce as many sets of Handbook forms (pgs. 66-80) as you have objectives to be validated.

Step 2: Order the major objectives (as provided in the "Application for Validation") from the most important to the least important. This should have team consensus.

Step 3: Take the objectives, one at a time starting with the most important, through the set of Effectiveness/Success questions (1, 2, 3, 4 (a-i), 5, 6, 7, pages 66-80) and assign a whole number value on the scale of 0 to 3 to each question for each objective.

Step 4: After assigning a scale value to each question for each objective, enter this scale value (except for Question 4) opposite the question number in the column under the appropriate objective number on Page 81 of the Handbook. Only whole numbers may be entered.

A. Special directions for completing and computing the final value for question #4:

1. Question 4(a) must receive a scale value greater than zero as per the instructions in the Handbook, Page 69. Enter this value in the "Evaluation" "Rating Points" Table at the bottom of Page 81 of the Handbook opposite question 4(a) in the column under the appropriate objective number. Again, only whole numbers may be entered.

2. Do the same for parts (b) through (i) of question 4 entering "NA" opposite 4(d) and/or 4(e) if appropriate.

3. Sum each column and enter that value in the row marked "Total" of the "Evaluation" "Rating Points" Table at the bottom of Page 81 of the Handbook.

4. Before you enter a whole number value for this "Evaluation" section beside question 4 on the "SUMMARY RATING" Table at the top of Page 81, you must first divide the "TOTAL" by either 7, 8 or 9. A table of values has been provided for you on Page of the SMA Guidebook to facilitate this process.

Step 5: When all scale values (questions 1 - 7) have been entered in the "SUMMARY RATING" Table for each objective, sum each column and enter these totals in the spaces opposite "SUMMED RATING POINTS" (below each column).

Note: In order for a project to be validated, the first 50% of its major objectives submitted for validation must each receive a "SUMMED RATING POINT" total equal to or greater than (≥) than 19. If a project submits an odd number of major objectives for validation, please refer to the table below to determine the minimum number of objectives that must have a SUMMED RATING POINT total equal to or greater than (≥) 19 in order to qualify for project validation.

# of objectives submitted	# of objectives needing 19 or greater SUMMED RATING POINTS
1	1
3	2
5	3
7	4
9	5
11	6
.	.
.	.
.	.
etc.	etc.

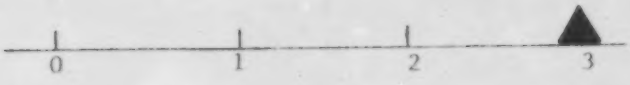
Step 6: Sum across the "SUMMED RATING POINTS" row, i.e., compute the total by adding the totals of each column.

Step 7: Divide the value arrived at in step 6 by the number of objectives submitted for validation (Round to the nearest whole number).

Step 8: In order for the Effectiveness/Success criterion to be validated the rounded value from Step 7 must be equal to or greater than (≥) 19. Enter this value on the "VALIDATION SUMMARY" sheet, Page 104 of the Handbook in the box opposite "Validation of Evidence on Effectiveness/Success".

OBJECTIVE 1

QUESTION 1



QUESTION 2

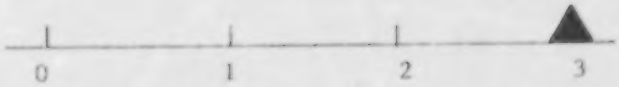
QUESTION 3

QUESTION 4(a)

4(b)

4(c)

4(d)



4(e)

4(f)

4(g)

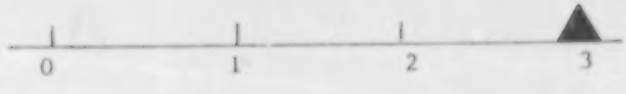
4(h)

4(i)

QUESTION 5

QUESTION 6

QUESTION 7



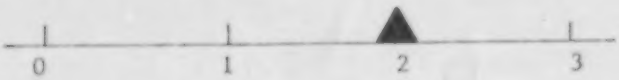
OBJECTIVE 2

QUESTION 1

QUESTION 2

QUESTION 3

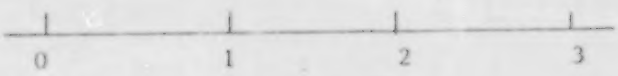
QUESTION 4(a)



4(b)

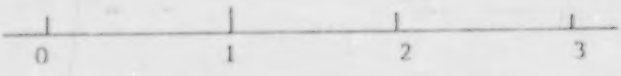
4(c)

4(d)



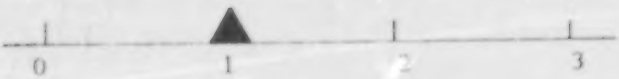
NA

4(e)



NA

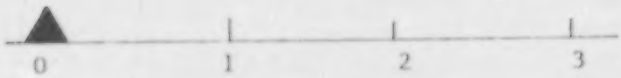
4(f)



4(g)

4(h)

4(i)



QUESTION 5

QUESTION 6

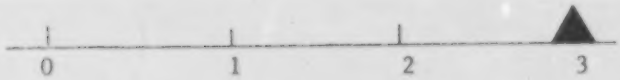
QUESTION 7

OBJECTIVE 3

QUESTION 1

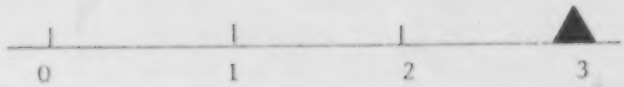
QUESTION 2

QUESTION 3



QUESTION 4(a)

4(b)



4(c)

4(d)

4(e)

NA

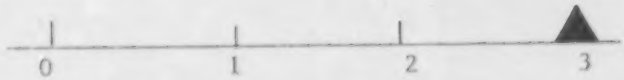
4(f)

4(g)

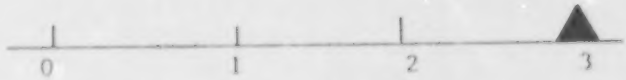
4(h)

4(i)

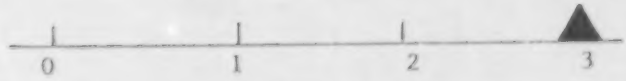
QUESTION 5



QUESTION 6



QUESTION 7



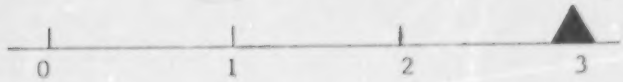
OBJECTIVE 4

QUESTION 1

QUESTION 2

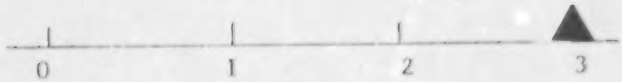
QUESTION 3

QUESTION 4(a)



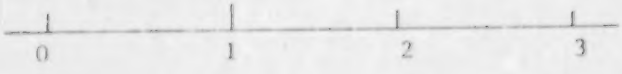
4(b)

4(c)



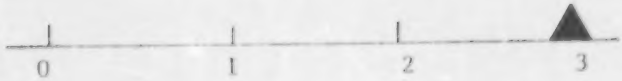
4(d)

4(e)



NA

4(f)

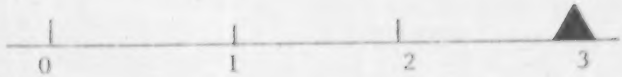


4(g)

4(h)

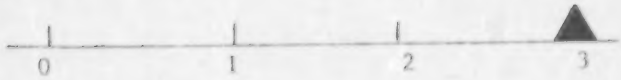
4(i)

QUESTION 5



QUESTION 6

QUESTION 7



OBJECTIVE 5

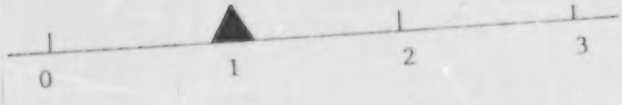
QUESTION 1

QUESTION 2

QUESTION 3

QUESTION 4(a)

4(b)



4(c)

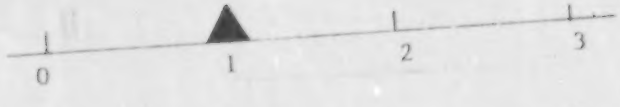
4(d)

NA

4(e)

4(f)

4(g)



4(h)

4(i)

QUESTION 5

QUESTION 6

QUESTION 7

PART II--Validation of Evidence on Effectiveness/Success

SUMMARY RATING

Evidence of Effectiveness/Success Items	Rating Points: Nominated Objectives by Number 1	2	3	4	5	6	7
*1. Measurability of objectivity							
2. Needs determination							
3. Intensity of project activities							
4. Evaluation (See directions below.)							
*5. Attainment of objective							
*6. Achievement and learner change and generalizability of project							
*7. Statistical and educational significance							
SUMMED RATING POINTS (Summed Rating Points Must Total 19 for Validation of Each Objective.)							
Total Scaled Score(Summed Score Divided by 7) Transfer Scaled Score to p. 104.							

Add additional objectives as needed.

*Items 1, 5, 6, and 7 must each receive a rating of 3 points. If any of these items does not receive a rating of 3 points, reject the objective from further validation.

Directions for Item 4 (Evaluation, parts a-i): Enter rating for each item. For each objective, total the rating (parts a-i). Divide the total by 9 (the number of subitems).
If item 4(d) or 4(e) is rated NA (not applicable), divide the total by 8.
If item 4(d) and 4(e) are rated NA (not applicable), divide the total by 7.
Enter result under item 4 in summary record above.

4. Evaluation	Rating Points: Nominated Objectives by Number 1	2	3	4	5	6	7
(a). Evaluation design							
(b). Evaluation procedures							
(c). Project activities							
(d). Sampling techniques							
(e). Control group selection							
(f). Instrumentation							
(g). Qualified personnel							
(h). Data accuracy							
(i). Data analysis procedures							
TOTAL							

"Evaluation" (4, a-i) TABLE OF VALUES *

DIRECTIONS: Record the value in parentheses ( ) opposite question #4 on the SUMMARY RATING sheet on page 81.

	Values obtained when 7, 8 or 9 are used as divisors as per Handbook instructions					
N	N/7	Record	N/8	Record	N/9	Record
3	.43	(0)	.38	(0)	.33	(0)
4	.57	(1)	.50	(1)	.44	(0)
5	.71	(1)	.63	(1)	.56	(1)
6	.86	(1)	.75	(1)	.67	(1)
7	1.00	(1)	.88	(1)	.78	(1)
8	1.14	(1)	1.00	(1)	.89	(1)
9	1.29	(1)	1.13	(1)	1.00	(1)
10	1.43	(1)	1.25	(1)	1.11	(1)
11	1.57	(2)	1.38	(1)	1.22	(1)
12	1.71	(2)	1.50	(2)	1.33	(1)
13	1.86	(2)	1.63	(2)	1.44	(1)
14	2.00	(2)	1.75	(2)	1.56	(2)
15	2.14	(2)	1.88	(2)	1.67	(2)
16	2.29	(2)	2.00	(2)	1.78	(2)
17	2.43	(2)	2.13	(2)	1.89	(2)
18	2.57	(3)	2.25	(2)	2.00	(2)
19	2.71	(3)	2.38	(2)	2.11	(2)
20	2.86	(3)	2.50	(3)	2.22	(2)
21	3.00	(3)	2.63	(3)	2.33	(2)
22	XXXX	XXX	2.75	(3)	2.44	(2)
23	XXXX	XXX	2.88	(3)	2.56	(3)
24	XXXX	XXX	3.00	(3)	2.67	(3)
25	XXXX	XXX	XXXX	XXX	2.78	(3)
26	XXXX	XXX	XXXX	XXX	2.89	(3)
27	XXXX	XXX	XXXX	XXX	3.00	(3)

Possible Point Variation in 4(a-i)

*All possible values derived from computations under "Directions for Item 4" on page 81. If one extrapolates downward from 3 to an N of 1 or 2 the recorded value would be 0.

THE LOGICAL STRUCTURE OF EVALUATION DESIGN IS THE SAME FOR ALL TYPES OF EVALUATION

The parts, briefly, are as follows:

A. FOCUSING THE EVALUATION
   1. Identify the major level(s) of decision-making to be served, e.g., local, state, or national.
   2. For each level of decision-making, project the decision situations to be served and describe each one in terms of its locus, focus, timing, and composition of alternatives.
   3. Define criteria for each decision situation by specifying variables for measurement and standards for use in the judgment of alternatives.
   4. Define policies within which the evaluation must operate.

B. COLLECTION OF INFORMATION
   1. Specify the source of the information to be collected.
   2. Specify the instruments and methods for collecting the needed information.
   3. Specify the sampling procedure to be employed.
   4. Specify the conditions and schedule for information collection.

C. ORGANIZATION OF INFORMATION
   1. Specify a format for the information which is to be collected.
   2. Specify a means for coding, organizing, storing, and retrieving information.

D. ANALYSIS OF INFORMATION
   1. Specify the analytical procedures to be employed.
   2. Specify a means for performing the analysis.

E. REPORTING OF INFORMATION
   1. Define the audiences for the evaluation reports.
   2. Specify means for providing information to the audiences.
   3. Specify the format for evaluation reports and/or reporting sessions.
   4. Schedule the reporting of information.

F. ADMINISTRATION OF THE EVALUATION
   1. Summarize the evaluation schedule.
   2. Define staff and resource requirements and plans for meeting these requirements.
   3. Specify means for meeting policy requirements for conduct of the evaluation.
   4. Evaluate the potential of the evaluation design for providing information which is valid, reliable, credible, timely, and pervasive.
   5. Specify and schedule means for periodic updating of the evaluation design.
   6. Provide a budget for the total evaluation program.

The utilization of this review process should positively identify strengths and weaknesses in any existing evaluation design.

CHECK LIST

MINIMUM ESSENTIALS FOR EVALUATION DESIGNS

	Missing	Poor	Fair	Good
I. General Program Design- - - - - - - - - - -				
A. Task-areas and strategy-dimensions are defined in relation to general objectives				
B. Range of activity-components is designed in relation to general objectives				
II. Product Outcomes (attainment of specific performance objectives)- - - - -				
A. Specific performance objectives indicate in detail-level and scope:				
(1) Nature of performance (behavior or material, as appropriate) expected of the target individual or group				
(2) Direction or level (if basis for prediction exists) of expected performance accomplishment				
(3) Primary conditions under which performance is expected to be conducted (when measured)				
(4) Units of performance measurement				
B. Measurement techniques are described				
C. Reliable and valid measurement instruments are selected				
D. Measurement-process design specifies:				
(1) Who, or what, is to be measured (if not in specific objectives)				
(2) Responsibility for measurement				
(3) When measurements are to be made (Schedule of complete cycle)				
(4) Conditions of measurement (if not in specific objectives)				
E. Data-collection procedures are described				

	Missing	Poor	Fair	Good
F. Data-analysis techniques are specified				
G. Design for presentation of data analysis includes:				
(1) Format				
(2) Reporting mechanisms				
(a) responsibility				
(b) procedures				
(c) recipients				
(d) schedule				
III. Operational Process (means of attaining specific objectives)				
A. Specific operational-process descriptions establish relationship to produce outcomes				
B. Specific operational-process descriptions indicate:				
(1) Nature of performance (behavior expected of the operator)				
(2) Direction or level (if basis for prediction exists) of expected performance accomplishment				
(3) Primary conditions under which performance is expected to be conducted (when measured)				
(4) Units of performance measurement				
C. Measurement techniques are described				
D. Reliable and valid measurement instruments are selected				
E. Measurement-process design specified:				
(1) Who, or what, is to be measured (if not in specific objectives)				
(2) Responsibility for measurement				
(3) When measurements are to be made (schedule of complete cycle)				
(4) Conditions for measurement (if not in specific objectives)				
F. Data-collection procedures are described				
G. Data-analysis techniques are specified				
H. Design for presentation of data analysis includes:				
(1) Format				
(2) Reporting mechanisms				
(a) responsibility				
(b) procedures				
(c) recipients				
(d) schedule				
IV Management Process (means of operational-process performance control)				
A. Specific management-process descriptions establish relationship to produce outcomes				

	Missing	Poor	Fair	Good
B. Specific management-process descriptions indicate:				
(1) Nature of performance (behavior expected of the manager				
(2) Direction or level (if basis for prediction exists) of expected performance accomplishment				
(3) Primary conditions under which performance is expected to be conducted (when measured)				
(4) Units of performance measurement				
C. Measurement techniques are described				
D. Reliable and valid measurement instruments are selected				
E. Measurement-process design specifies:				
(1) Who, or what, is to be measured (if not in specific objectives)				
(2) Responsibility for measurement				
(3) When measurements are to be made (schedule of complete cycle)				
(4) Conditions for measurement (if not in specific objectives)				
F. Data-collection procedures are described				
G. Data-analysis techniques are specified				
H. Design for presentation of data analysis includes:				
(1) Format				
(2) Reporting mechanisms				
(a) responsibility				
(b) procedures				
(c) recipients				
(d) schedule				

Goal #1 INCREASE THE EDUCATIONAL OPPORTUNITIES FOR PRESCHOOL CHILDREN

WORK BREAKDOWN STRUCTURE

Objectives	Activities	Work Packages	Tasks
The number of preschool children enrolled in programs in Title I areas will be increased by 10% by expanding existing programs and in the initiation of new programs during the 1971-72 school year.	Disseminate information of successful N.J. Title I Preschool programs.	Determine successful programs.	Locate preschool programs. Visit programs. Gather objective and subjective data. Determine degree of success. Select program to disseminate.
		Gather information to disseminate.	Conduct interviews. Photograph activities. Organize data to show progress or growth.
		Develop dissemination project.	Determine audience. Determine medium. Select format. Develop message. Develop visuals. Organize. Print.
		Disseminate project information.	Schedule distribution. Distribute.
		Evaluate results.	Determine format. Develop instrument. Pilot instrument. Compile results. Assess results.
	Conduct six Regional meetings to encourage development of Title I preschool programs.	Plan content and format.	Determine content. Select consultants. Schedule consultants. Develop materials.

Objectives	Activities	Work Packages	Tasks
		Arrange location and dates.	Select location. Finalize logistics.
		Invite participants.	Determine participants. Write brochure. Mail brochure. Receive replies.
		Conduct meeting.	Deliver materials. Set up facilities. Conduct meeting.
		Evaluate results.	Plan format. Develop instrument. Conduct evaluation. Compile results.
	Place emphasis on preschool programs in the Title I Guidelines and proposal format.	Conduct review of existing Title I Proposal.	Review problem areas at State level. Review problems with Title I coordinators. Determine additional needs. Compile results.
		Develop 1971-72 proposal format.	Determine information required in new proposals. Determine method to place emphasis on preschool program. Develop 1971-72 format. Pilot format. Revise or adapt. Print format.

Objectives	Activities	Work Packages	Tasks
		Design guidelines to accompany Title I proposals.	Review existing guidelines and proposal. Review results of proposal pilot. Write new guidelines. Pilot guidelines. Print guidelines.
		Disseminate proposals and guidelines.	Determine training needs. Schedule meeting. Schedule personnel. Invite participants. Deliver materials. Conduct meeting.
	Disseminate results of research in the areas of preschool education.	Conduct search of research activities.	Determine sources of research information. Gather results of research activities. Evaluate results. Determine information which is relevant to N.J. and Title I.
		Plan Dissemination Activities.	Determine audience. Determine format. Develop format. Disseminate.
		Evaluate results.	Determine format. Develop instrument. Conduct evaluation. Tabulate results. Review and report.

Goal #1 / Activities	Objective S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J
1. Disseminate information of successful N. J. Title I preschool programs.	o	o	o	△													
2. Conduct six Regional meetings to encourage development of Title I preschool programs.	o	o	△														
3. Place emphasis on preschool programs in the Title I Guidelines and Proposal format.	o	o	△														
4. Disseminate results of research in the areas of preschool education.	o	o	△														

Goal #1 / Work Packages	Objective S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J
1. Determine successful programs.	△																
2. Gather dissemination information.		△															
3. Develop dissemination project.			△														
4. Disseminate project information.				△													
5. Evaluate results.					△												
6. Plan content and format.						△											
7. Arrange location and dates.							△										
8. Invite participants.								△									
9. Conduct meeting.									△								
10. Evaluate results.										△							
11. Conduct review of existing Title I prop.											△						
12. Develop 1971-72 proposal format.												△					
13. Design guidelines to accompany Title I proposals.													△				
14. Disseminate proposals and guidelines.														△			
15. Conduct search of research activities.															△		
16. Plan dissemination activities.																△	
17. Evaluate results.																	△

Goal #1 / Tasks	Objective S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J
1. Locate preschool programs.																	
2. Visit programs.																	
3. Gather objective & subjective data.																	
4. Determine degree of success.																	
5. Select program to disseminate.																	
6. Conduct interviews.																	
7. Photograph activities.																	
8. Organize data to show progress or growth.																	
9. Determine audience.																	
10. Determine medium.																	
11. Select format.																	
12. Develop message.																	
13. Develop visuals.																	
14. Organize.																	
15. Print.																	
16. Schedule distribution																	
17. Distribute.																	
18. Determine format.																	
19. Develop instrument.																	
20. Pilot instrument.																	
21. Compile results.																	
22. Assess results.																	
23. Determine content.																	
24. Select consultants.																	
25. Schedule consultants.																	
26. Develop materials.																	
27. Select location.																	
28. Finalize logistics.																	
29. Determine participants.																	
30. Write brochure.																	
31. Mail brochure.																	
32. Receive replies.																	
33. Deliver materials.																	

Goal #1 / Tasks	Objective S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J
34. Set up facilities.																	
35. Conduct meeting.																	
36. Plan format.																	
37. Develop instrument.																	
38. Conduct evaluation.																	
39. Compile results.																	
40. Review problem areas at State level.																	
41. Review problems with Title I coordinators.																	
42. Determine additional needs.																	
43. Compile results.																	
44. Determine information required in new proposals.																	
45. Determine method to place emphasis on preschool program.																	
46. Develop 1971-72 format.																	
47. Pilot format.																	
48. Revise or adapt.																	
49. Print format.																	
50. Review existing guidelines & prop.																	
51. Review results of proposal pilot.																	
52. Write new guidelines.																	
53. Pilot guidelines.																	
54. Print guidelines.																	
55. Determine training needs.																	
56. Schedule meeting.																	
57. Schedule personnel.																	
58. Invite participants.																	
59. Deliver materials.																	
60. Conduct meeting.																	

Goal #1 / Tasks	Objective S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J
61. Determine sources of research information.																	
62. Gather results of research activities.																	
63. Evaluate results.																	
64. Determine information which is relevant to N.J. & Title I.																	
65. Determine audience.																	
66. Determine format.																	
67. Develop format.																	
68. Disseminate.																	
69. Determine format.																	
70. Develop instrument.																	
71. Conduct evaluation.																	
72. Tabulate results.																	
73. Review & report.																	

Control Group Assumptions

In any design that involves comparing two or more groups of subjects who have been exposed to different experimental treatments, there is an underlying assumption that the groups being compared are equivalent before the introduction of the treatments. However, the investigator cannot simply make this assumption; he must take steps to see that it is met.

Clearly, the task of creating or of unearthing groups that are equivalent in all respects is an impossible one. Before considering how one gets around this problem, it is necessary to distinguish two different reasons for wanting equivalent experimental and control groups. The first is to provide a basis for inferring that differences which may be found on the dependent variable do not result from initial differences between the two groups, in terms either of position on the dependent variable or of other factors. The goal here is to ensure, as far as possible, the validity of the inferences made on the basis of the experiment. But there is a second goal, that of increasing the sensitivity of the experiment -- i.e., increasing its ability to register small effects of the experimental treatment that might be obscured by the effects of other factors.

These two goals call for somewhat different procedures in establishing the equivalence of groups. The goal of protecting the validity of the experiment by ensuring that experimental and control groups differ initially only by chance is achieved by procedures termed randomization. The goal of increasing the sensitivity of the experiment, so that the effects of the causal variable will be apparent even if they are relatively small or if there are relatively few subjects, is achieved by matching procedures.

Randomization. Randomization provides the basic safeguard against differences between experimental and control groups that might lessen the validity of inferences about the effects of the experimental treatment.

Matching. Although random assignment, where it is feasible, is generally considered to provide adequate protection against interpreting differences on the dependent variable as resulting from the independent variable when in fact they stem from prior differences between the two groups, it is not the most effective procedure from the point of view of increasing the sensitivity of the experiment. In the interest of research efficiency, it is desirable that the experiment reveal true differences brought about by the experimental treatment, even if they are small in relation to differences produced by other variables. Matching is not a substitute for random assignment, but a supplement to it. Matching procedures can take account of only a few variables; those that are unaccounted for should be randomly distributed between the experimental and the control groups.

FACTORS JEOPARDIZING INTERNAL AND EXTERNAL VALIDITY

Fundamental to the following listing is a distinction between internal validity and external validity. Internal validity is the basic minimum without which any experiment is uninterpretable: Did in fact the experimental treatments make a difference in this specific experimental instance? External validity asks the question of generalizability: To what populations, settings, treatment variables, and measurement variables can this effect be generalized?

Relevant to internal validity, eight different classes of extraneous variables will be presented; these variables, if not controlled in the experimental design, might produce effects confounded with the effect of the experimental stimulus. They represent the effects of:

1. History, the specific events occurring between the first and second measurements in addition to the experimental variable.
2. Maturation, processes within the respondents operating as a function of the passage of time per se (not specific to the particular events), including growing older, growing hungrier, growing more tired, and the like.
3. Testing, the effects of taking a test upon the scores of a second testing.
4. Instrumentation, in which changes in the calibration of a measuring instrument or changes in the observers or scores used may produce changes in the obtained measurements.
5. Statistical regression, operating where groups have been selected on the basis of their extreme scores.
6. Biases resulting in differential selection of respondents for the comparison groups.
7. Experimental mortality, or differential loss of respondents from the comparison groups.
8. Selection-maturation interaction is confounded with, i.e., might be mistaken for, the effect of the experimental variable.

The factors jeopardizing external validity or representativeness which will be discussed are:

9. The reactive or interaction effect of testing, in which a pretest might increase or decrease the respondent's sensitivity or responsiveness to the experimental variable and thus make the results obtained for a pretested population unrepresentative of the effects of the experimental variable for the unpretested universe from which the experimental respondents were selected.
10. The interaction effects of selection biases and the experimental variable.
11. Reactive effects of experimental arrangements, which would preclude generalization about the effect of the experimental variable upon persons being exposed to it in nonexperimental settings.
12. Multiple-treatment interference, likely to occur whenever multiple treatments are applied to the same respondents, because the effects of prior treatments are not usually erasable.

TABLE 1

SOURCES OF INVALIDITY FOR DESIGNS 1 THROUGH 6

	Sources of Invalidity											
	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Pre-Experimental Designs:												
1. One-Shot Case Study X O	−	−				−	−			−		
2. One-Group Pretest-Posttest Design O X O	−	−	−	−	?	+	+	−	−	−	?	
3. Static-Group Comparison X O --- O	+	?	+	+	+	−	−	−		−		
True Experimental Designs:												
4. Pretest-Posttest Control Group Design R O X O R O O	+	+	+	+	+	+	+	+	−	?	?	
5. Solomon Four-Group Design R O X O R O O R X O R O	+	+	+	+	+	+	+	+	+	?	?	
6. Posttest-Only Control Group Design R X O R O	+	+	+	+	+	+	+	+	+	?	?	

Note: In the tables, a minus indicates a definite weakness, a plus indicates that the factor is controlled, a question mark indicates a possible source of concern, and a blank indicates that the factor is not relevant.

It is with extreme reluctance that these summary tables are presented because they are apt to be "too helpful," and to be depended upon in place of the more complex and qualified presentation in the text. No + or − indicator should be respected unless the reader comprehends why it is placed there. In particular, it is against the spirit of this presentation to create uncomprehended fears of, or confidence in, specific designs.

BASIC INTERVIEW TECHNIQUES

A. The Fundamentals

In this section, the validator will learn techniques which are fundamental to conducting a good validation interview.

More advanced interviewing techniques are covered in the last section of this guideline, but it is most important that the validator master the fundamentals. He must learn how to avoid major interviewing errors (give-away questions, questions which yield little information, questions which put the interviewee on the defensive). Later will come additional techniques: how to construct problem questions, how to use reassurance, and how to probe what the interviewee says.

1. Correcting Common Errors

   This is how one validator tried to find out why a member of the Advisory Committee wanted to see a specific project change its erphasis.

   Validator: Do you like the present program?
   Interviewee : It is okay.

   The validator did not find out what he wanted to know. The question can be answered in one word. It was not constructed so as to elicit an informative answer.

2. Questions Yielding Little or No Information

   A very common error among validators is the use of questions which add little or no information about the areas in which they have interest. Frequently, these can be answered in one word. For example:

   a. Do you like the present program?
   b. Do you feel that the Board of Education has reached a dead-end in policy-making matters?
   c. Does this type of organization structure appeal to you more than the original structure?

   We are not advocating that the validator use only questions which yield lengthy answers. One word, given in response to your question, may give exactly the information you want. For example, if the validator wanted to find out how many aides were in a program for a specific period, he could ask the question this way:

   (1) How many students were enrolled in this program last year?
   (2) How many students completed the training for this program last year?

3. Give-Away Questions

   Both novice and experienced validators commit this error frequently. They ask questions which "give away" the "right" answer, that is, the answer that the validator thinks the interviewee should give. The following is taken from an actual interview conducted by an experienced validator:

   Validator: The answer that you gave us regarding the way in which program priorities are determined seems to be in conflict with answers

supplied by the Principal. Are you at odds with each other, or were you nervous?

Interviewee: I was nervous.

The validator gave the interviewee the answer as well as the question.

Following are examples of other "give-away" questions:

a. I assume that there is no problem of coordination and supervision, right?
b. You'd recommend hiring additional staff to handle this problem, wouldn't you?

4. Antagonizing Questions

In addition to give-away questions, which add little to their knowledge, validators frequently use antagonizing questions that put the interviewee on the defensive. The following exchange illustrates an error of this type:

Validator: According to our records, you change jobs every year or so. Why do you do so much job hopping?

Interviewee: (Angrily) -- As a matter of fact, I don't happen to be a job hopper....

By his choice of words, the validator antagonized the interviewee and put him on the defensive. If the validator puts the interviewee on the defensive, he is likely to clam up or lie; whereas if he is put at ease, he is more likely to speak freely and tell the truth.

There are three ways to avoid antagonizing questions:

a. Use neutral rather than emotionally loaded words;
b. Use softening introductory phrases;
c. Use qualifiers.

5. Use of Neutral Words, Introductory Phrases and Qualifiers

Some words are loaded with emotional overtones. The validator should try to choose words which are neutral or "unloaded".

Some examples are:

a. disagreed
b. unsatisfactory
c. dislike
d. lack of skill
e. informal
f. terminate
g. frequent change
h. leave

6. Introductory Phrases

A good validator will usually ask many questions. This often makes the interviewee feel as if he is being grilled or cross-examined and so tends to make him defensive. To make questions sound more conversational, and less like an interrogation, the validator should "soften" them with <u>introductory phrases</u>.

For example:

a. To what do you attribute the problems encountered in this program?
b. Would you say that the board is representative of the ethnic composition of the neighborhood?
c. Is it possible that aides are not being trained in marketable skills?
d. How did you happen to choose your Personnel Director?

e. What prompted the choice of this program over other programs?

7. Use of Qualifiers
Finally, an interviewee will tend to become defensive if the validator poses questions in terms of black and white, rather than varying shades of grey. A validator will be able to elicit more information by using qualifying words which introduce the notion of degree. For example: a bit, to some extent, partially.
This has a tendency to "soften" or remove the sting from a question:
a. Were you somewhat dissatisfied with your organizational structure?
b. Did you tend to hamper program implementation?
c. Do you get slightly irritated when your executive staff doesn't follow your instructions?

B. Constructive Techniques
This section is concerned with more constructive techniques used to facilitate the validation interview. These will help the validator to learn how an interviewee would act in certain on-the-job situations, to elicit information which the interviewee may be hesitant to reveal and to follow up leads in the interviewee's response. The techniques are: using problem questions, using reassurance, and probing.

1. Problem Questions
If a validator wanted to find out how an interviewee would react in certain situations which are likely to occur in an on-the-job situation, one of the easiest ways to get at this is to simply describe the situation and ask him how he would handle it. This is known as a problem question. The easiest way to construct a problem question is to begin with "what would you do if..." and then present the situation.
For example:
a. What would you do if a staff member were fired and asked for a grievance hearing?
b. Suppose one of your programs was not being effectively administered?
c. How would you handle the situation in which one of the Advisory Board members was always trying to monopolize the Board meetings?
The validator should try, if possible, to make the interviewee feel that he does not blame him for the difficulties.

Interviewee: They funded a program with no guidelines and then got angry when we did it our way.

Validator: So, through no fault of yours, the program is not up to par. That happens all the time.

Interviewee: Then they merged the Consumer Education program with the Community Action Program and cut back the staff.

Validator: This often happens after a merger. Programs are often merged so that personnel required to operate them can be cut back.

2. Using Reassurance
Persons being interviewed, as you know, are not likely to reveal unfavorable information about programs, and yet that is the kind of information that the validator will often want. The validator will usually have to rely on former employees, former students, and his own observations and inferences for such information. If, however, an interviewee seems on the verge of revealing such information, the validator can often obtain it by reassuring him. An interviewee who has begun to reveal such information can be encouraged to continue by your reassurance. For example, when an interviewee is relating a problem or a difficult situation to you, you can reassure him by indicating that you realize that this is indeed a problem or difficulty.

Interviewee: So, OE calls us on the carpet for something that was really its fault.

Validator: That really is a problem; that's not a very pleasant position to be in.

Another way to reassure an interviewee is to point out that his problem or error is a very common one.

3. Probing
As a validator, you will often discern levels at which the interviewee hesitates and will want to follow these up by probing into certain details.
For example:

Interviewee: Sometimes we would really have a lot of people at a community meeting and get things accomplished.

The word "sometimes" suggests that the meetings were not always well attended and that things were not always accomplished at these meetings.

Validator: Tell me more about these meetings. What sort of meeting should the community have? Approximately how many of the community meetings would you say were really good?

ADVANCED INTERVIEW TECHNIQUES

A. Additional Techniques
The previous sections have been primarily concerned with eliciting information through direct questioning techniques. These techniques will be adequate for most interviews. However, as he gains experience, the validator may find it desirable to increase his skills by perfecting some additional techniques. This section will be concerned with interviewing techniques which will improve the validator's skill in indirect questioning and increase his ability to draw out an interviewee.

1. Self Evaluation
Interviewees are not often the best judges of their own programs. They often rate the effectiveness of their programs higher than they would be rated by participant observers and non-participants. Suppose, as a validator,

you are interviewing a project director and trying to elicit information about coordination and supervision. You could of course ask him a direct question about this skill, but you are more likely to get reliable information if you ask the employees whom he supervises in the various programs. Questions requiring an answer based on subjective self-evaluation have limited value. In order to obtain objective information, it is usually necessary to ask indirect questions. Indirect questions are those which elicit information that can be used to make inferences about validity and reliability. In validating a program, it is often important for the validator to try to determine the interest, attitudes and motivations of persons involved in the program. This is sometimes difficult to do because interviewees naturally tend to give answers which they believe will be preferred by the validator.

In previous sections, the validator learned to avoid give-away questions in which the phrasing of the question indicated the "correct" answer, but rephrasing by itself may not remove the give-away quality. The context frequently gives the interviewee the clue to the "correct" answer. There are two general approaches to getting information about interests and attitudes without suggesting the answer which would be most favorable. One is by being indirect in questioning and the other is by drawing out the interviewee.

2. Indirect Questioning
   It is often necessary to get information indirectly when it is unlikely that this information can be accurately elicited directly. The key word is inference. The validator must infer something about a program from what the interviewee says in talking about the program. The validator should try to base his inferences on several lines of evidence. If he makes his inference on the basis of one remark or one incident, he may make a serious error. The only categorical statement which can be made about when to use indirection is this: Use it whenever you cannot be sure that direct questioning will yield accurate and objective information. The validator should, however, bear in mind that some interviewees will be more honest than others, both with themselves and with him.

3. Laundry List
   One advantage of using a laundry list of questions is that you can present a number of alternatives from which an interviewee cannot choose because none of them seems reasonable to him. This forces him to clarify his position. Answers to a laundry list of questions can tell the validator something about interest and attitudes, but its use sometimes restricts the range of possible answers by offering a limited set of alternatives.

4. Open-Ended Questions
   A validator should realize the extent to which his question

structures or determines the answer he gets. At one extreme, there are questions which give the interviewee a wide range of possible answers. These are called open-ended. At the other, there are questions which restrict him to one or several possible answers. These are called structured. For example:

a. Open-Ended: Where do you see this project going in the next five years?

b. Structured: Which would you prefer-- close coordination and supervision, or relaxed coordination and supervision?

Frequently, a validator will use an open-ended question to "open-up" an interviewee in the area of interest. He then can follow up with the more specific questions based upon what the interviewee says. An open-ended question will also give the validator a chance to see what an interviewee regards as important in a given area-- as opposed to what he himself thinks important.

5. Reflection

Sometimes an interviewee will need another prod to keep him talking, even after the validator has opened him up, particularly on sensitive topics. One good technique is to reflect what he says more or less like a mirror:

Interviewee: When we first got our funds for the program, we had a hard time recruiting staff.

Validator: You had a hard time recruiting staff?

A reflection is a simple restatement of all or a portion of what an interviewee says. It is not a question of probing of what he says.

6. Interpretation

A technique closely related to reflection is interpretation; but whereas reflection involves the simple restatement of what an interviewee says, interpretation is an attempt to state the meaning of what is said.

Interviewee: If we didn't have the U.S. Office of Education or the State Department of Education watching every move that we make, we could do a much better job.

Validator: You feel that you could do a better job when you are not closely supervised?

The technique of interpretation is often necessary for you, as a validator, to be sure that you understand what the interviewee means, particularly when he gives you a lengthy or disorganized answer:

Interviewee: Well, we didn't really like the idea of a program like that, but then we didn't have any program at all, and the students in the school could use the services, and also it would provide some additional jobs. So we thought that we would accept the program. If it led to something, fine, and if it didn't, at least the students and their parents knew we were trying and would not have been demonstrating or rioting, like they did in Watts.

Validator: So you accepted the program as a stop gap measure until a more substantial, long-range program could be developed?

Sometimes you may not really understand what the interviewee means, and may interpret what he says inaccurately. If your interpretation is accurate, the interviewee will probably agree with it or elaborate upon it. If it is not accurate, he will most likely clarify it. Sometimes you may want to give an inaccurate interpretation deliberately to "force" the interviewee to clarify his statement:

Validator: So, the only reason you accepted this program was to prevent riots.

Interviewee: No, that's not it at all... We accepted because...

INDICES OF SPECIFIC COMPONENTS OF TITLE III PROJECTS

A. Advisory Committee
   1. Its elections, selection responsibilities and functions
      a. As given in charter, constitution or by-laws
      b. As practiced
   2. Its committees, and their operation
      a. Number and types
      b. What do they do?
   3. Its representativeness of the general community
      a. Size
      b. Tenure of membership of each member
      c. Frequency of attendance of each member
      d. Members' knowledge of project, its aims and activities
      e. Sex, age, education, occupation of members
      f. Its relation with other agencies in the area
   4. Its meetings
      a. Frequency and length
      b. Activities engaged in
      c. Matters discussed and considered
      d. Time and effort spent on each item
   5. Its role
      a. Formulation of policy
      b. Determination of policy
      c. Execution of policy
      d. Program planning
      e. Advisory

B. Program or Service
   1. Its goals and objectives
      a. Services or program offered and planned in work statement
      b. Current status of services and programs proposed in work plan

2. Its administration and coordination
   a. Integration of components
   b. Supervision of activities
   c. Coordination with other agencies

3. Its effectiveness
   a. Number of people served, and how
   b. Location and hours
   c. Success in ameliorating a condition of educational deficiency
   d. Physical facilities and equipment
   e. Ability to involve neighborhood groups or other agencies

4. Development of trade-off models
   a. Summer programs
   b. Special programs to meet local educational needs

C. Professional Staff
1. Its size and calibre
   a. As given in the original proposal
   b. Outlined in job specifications
   c. In-service training
   d. Relevant experience
   e. Formal education

2. Ability to work with other staff members and members of neighborhood groups
   a. Giving technical assistance
   b. Attitude toward program and community groups

3. Administration
   a. Coordination
   b. Supervisory ability (what is this?)

4. Carrying out and developing programs
   a. Attitude toward program
   b. Innovation within program

D. Organization Structure
1. Its communication system
   a. Up and down (feedback)
   b. Formal and informal
   c. Does it help or hinder?

2. Power flow
   a. Inside and outside
   b. Who makes the decisions?

3. Levels of supervision
   a. Number and types
   b. What do they do?

4. Reporting system
   a. Frequency
   b. Type

UNOBTRUSIVE MEASURES

The number of unobtrusive measures of learner affective behavior is limited only by the imagination of the validator. Some suggested unobtrusive measures include:

1. Frequency and type of extracurricular activities.
2. Frequency and type of elected positions.
3. Leisure activities.
4. Awards, citations, honors.
5. Number of books and periodicals read.
6. Peer group associations and participation.
7. Socially undesirable incidents such as intoxication, dope addiction, arrests, sexual deviation.
8. Referrals to counselor, psychologist, reading specialists, school nurse, etc.
9. Frequency of tardiness.
10. Grade point average as related to measured aptitude.
11. Grade placement as related to age.
12. Frequency and type of academic courses chosen.
13. Completion and qualitative judgments of homework assignments.
14. Number of dropouts.
15. Frequency and type of disciplinary actions taken.
16. Interpersonal relationships with both students and teachers.
17. Anecdotal records on student or teacher behavior relative to the attribute of interest.
18. Attendance at optional activities.
19. Frequency of student or teacher requests for changes in program or relationship (new teachers or students transferred out).
20. Vocational, avocational or educational choices expressed or carried out.

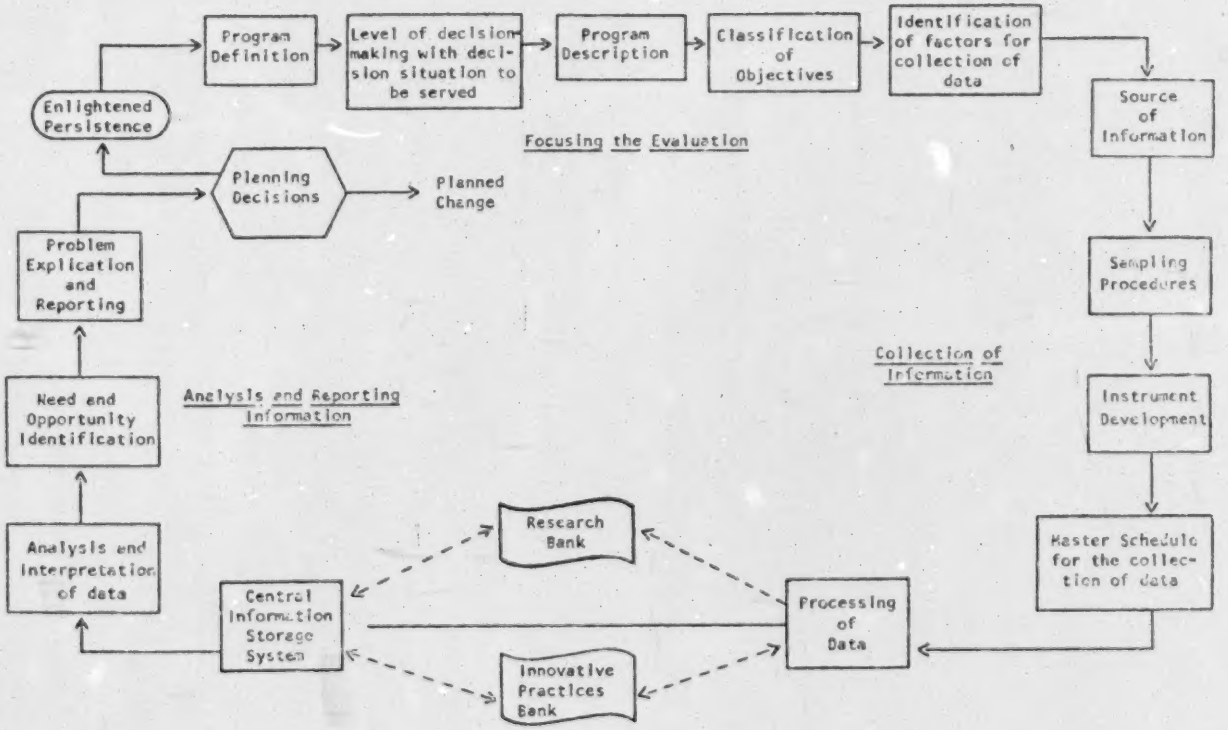
21. Student or teacher absences.

22. The frequency with which appointments are kept or broken.

23. Frequency of student or teacher publishing.

Many of the above listed measures can be taken directly from school records. Others require the development of checklists or rating scales.

Strategy for Context Evaluation



V. DISSEMINATION AND UTILIZATION STRATEGIES

The preceding review of specific one-way and two-way media of the knowledge transmission process has included suggestions for the most profitable use of each medium in a total dissemination strategy. In order to apply more adequately the general recommendations mentioned in the discussions of the individual media we now propose to relate these media directly to an overall dissemination and utilization strategy. Our goal is to illustrate the part that each medium can play in a progressing plan of knowledge utilization.

A. PROCESSES OF DISSEMINATION AND UTILIZATION

Chapter Two of this report introduced three perspectives of the dissemination and utilization process: problem-solving (P-S); social interaction (S-I); and research, development, and diffusion (R,D&D). (Chapter Ten and Eleven discuss these views in greater detail.)

1. The Problem-Solving Perspective

This view of the dissemination and utilization process stresses the ultimate user of the innovation. It assumes that utilization is instigated by a *need* within the user and proceeds for the purpose of *satisfying that need*. In the process of need satisfaction the user goes through the following activities (usually with some outside assistance): translation of need into a problem statement, diagnosis of the problem, search and retrieval of information that will be helpful for making a selection of the innovation, adaptation of the innovation to his own situation, trial of the innovation, and evaluation of the effectiveness of the trial in satisfying the original need.

2. The Social Interaction Perspective

This second perspective on the dissemination and utilization process focuses on the informal communications environment of the user, as seen by his position in the *network of social relations* in the group(s) in which he is a member. Viewing the process from the S-I perspective, the stages that each member will sooner or later pass through in the process of innovating are: awareness, interest, evaluation, trial, and adoption.

3. The Research, Development, and Diffusion Perspective

This perspective is based on the assumption of a *rational sequence* of phases by which an innovation is invented or discovered, developed, produced, and, finally, disseminated to the user. It is the only one of these three perspectives which does not approach innovation from the point of view of the user; in fact, it presumes that the user be fairly passive, though not irrational.

Each of the three perspectives is a valid representation of knowledge dissemination and utilization which is being carried on today. Their conceptualizations of the stages in the dissemination and utilization process differ. (See Chapter Ten) Each is appropriate for certain kinds of innovations and for certain types of user systems.

BEST COPY AVAILABLE

TABLE 9.1* Potential Media Uses in D&U Strategies

This listing is intended to be merely suggestive of the relationship of media to parts of the D&U process.

TRANSMISSION MEDIA		D&U STRATEGIES		
		R,D&D Processes (going on in the resource system)	P-S Processes (from the point of view of change agent & user system)	S-I Processes (going on in the user system)
One-Way Diffusion Transmissions	written word Oral Presentation Television & Radio Film Demonstration	Dissemination (might be used in combination in a multi-media marketing program)	Search and Retrieval of potential solutions	Awareness by all, awareness and interest by some opinion leaders, awareness, interest and evaluation by innovators
One-Way Feedback Transmissions	Public Archives Private Records Surveys/Polls Observation	Research and Development (problem & need assessment, market analysis, product testing & evaluation)	Diagnosis of problem and evaluation of the innovation	These transmissions rarely discussed by S-I theorists. Presumably they create a general readiness for considering new innovations.
	Referenda/Elections Petitions Letters Protests, Riots, Revolts	Impetus for new research & R&D efforts (through foundation & Federal support; movement of researchers into "fashionable" topic areas)		
Two-Way Transmissions	Dyadic Exchange Small Group Discussion	May play some role in various processes of R,D&D usually unspecified	Potentially useful for all stages: Translation Diagnosis Search & Retrieval Adoption Trial Evaluation	Vital for evaluation and decision to try-out and to adopt
	Large Group/Temporary System (e.g.'s: action research, collaborative action inquiry, organizational survey feedback, organizational "grid", training labs, derivation conference)	Dissemination		

BEST COPY AVAILABLE

1. The t-test

The t-test is a statistical model that is designed to investigate several types of questions. The one-sample t-test determines, within specified levels of probability, whether or not the population from which the sample was drawn has a given mean. The two-sample t-test determines, within specified limits of probability, whether or not the populations from which the samples were drawn have the same mean. Different models are used, depending upon the type of sample data under analysis. In yielding a probability statement of the differences between group means, the model considers mean differences, sample variability, sample size, and whether the data are correlated or independent.

2. The Analysis of Variance

When two or more groups or samples are available, the analysis of variance is a model used to test for differences, within specified limits of probability, between the means of those groups. The procedure enables one to analyze variances in such a manner that conclusions can be drawn about means. In the one-way situation (only one independent variable), the procedure is merely an extension of the t-test to situations where there are more than two groups.

Higher order analysis of variance is used to ascertain, within specified limits of probability, the effects of two or more independent variables on a dependent variable. Utilizing higher order analysis of variance, the researcher can test main effects and interaction effects. In higher order analysis of variance, the researcher must concern himself with whether he is dealing with a fixed, mixed, or random model, since the type of model used will, in the presence of a significant interaction effect, make a difference in the error term used. Among the types of designs available with this technique are factorial designs, repeated measures designs, Latin Square designs, and numerous permutations and combinations of the various techniques.

3. The Analysis of Covariance

When two or more groups or samples are available, the analysis of covariance is a model used to test for differences, within specified levels of probability, among the means of those groups, after initial compensation for differences among groups with respect to one or more control variables. Analysis of covariance is not a method to adjust for lack of random sampling; it is rather a technique that can be used to increase the precision of one's experiment if the control variable(s) selected is highly linearly correlated with the dependent variable.

Analysis of covariance can be used in simple-classification or multiple-classification forms and, when the assumptions are met, is a more powerful technique than analysis of variance because it will provide a reduced error term.

4. Correlational Analysis

When the words correlational analysis appear in the Metric-Metric cells of the taxonomy, they refer to the use of the Pearson product-moment correlation coefficient. Basically, correlational analysis provides the researcher with procedures for quantifying the measured relationships between two or more variables. The size of the product-moment correlation coefficient varies from -1.00 to +1.00, thus providing the evaluator with an estimate of the size and direction of a given relationship.

If the evaluator is interested in studying the relationship between two variables he could use the product-moment correlation coefficient. If he is interested in the relationship between one variable and a combination of two or more other variables considered simultaneously, he could use a multiple product-moment correlation coefficient. Multiple correlation provides an index of the relationship between a single metric variable and a composite.

It is often evident that a relationship of some interest may be explained, at least to some extent, on the basis of correlations with a third variable or composite. In such a case, the evaluator may wish to determine the degree to which the two variables of primary interest are related beyond the relation implied by correlations with a third variable. The partial correlation between the variables of primary interest provides a measure of their relationship independent of some other variate.

Two major concerns should be kept in mind when dealing with the types of statistics outlined above. First, the researcher should assure himself that there is a linear relationship among the variables being considered. If this linear relationship doesnot exist, adjustments should be made for lack of linearity. Second, it should be remembered that correlation does not necessarily imply causation. It usually requires carefully designed longitudinal correlational studies before one can legitimately begin to infer causation from correlational data. Procedures are available for determining the statistical significance of product-moment correlation coefficients.

Regression analysis, to be discussed in the next section, is closely related to the product-moment correlation coefficient. The major distinction between correlational and regression analysis is that, in correlational analysis, both of the measures involved are random variables, whereas in regression analysis, only one measure is variable while the other is given.

5. Regression Analysis

Regression analysis is usally presented in the educational and psychological literature as a model for making prediction on a given criterion from one or a set of predictors. The yield of the approach is a simple or multiple correlation coefficient and a regression equation composed of a set of weights that can be used to optimize prediction. The simple or multiple correction coefficient and the weights can be subjected to significance tests, resulting in decisions relative to the effectiveness of prediction and to which predictors do or do not contribute significantly to predictors, including all possible regressions, the backward elimination procedure, the forward selection procedure, and a number of variants on the aforementioned procedures.

The previously discussed analysis of variance and analysis of covariance models, on the other hand, are generally presented as methods for analyzing controlled experiments in which different groups are subjected to different treatments or treatment configurations. Conclusions are drawn, within specified limits of probability, in terms of the significance of differences in sets of means or mean differences.

In actuality, linear regression analysis, the analysis of variance, and the analysis of covariance are identical systems, all deriving from the general linear model. In practice, however, there are some differences. Generally, the textbooks in educational and psychological statistics treat the techniques separately, presenting wholly different algorithms. Regression requires the computation and inversion of the matrix of correlations among a group of independent variables, a great deal of computation for even a small number of independent variables. Classical analysis of variance, on the other hand, capitalizes on the mutual orthogonality of main effects and interactions to reduce substantially the computation required. However, with the availability of electronic computer facilities, this is generally no longer a consideration.

What is being stated here is that regression analysis can be used to investigate the same kinds of questions as the analysis of variance and covariance. In addition, the technique offers much greater flexibility than either classical analysis of variance or covariance. Its use need not be limited to predictive studies, as is often the case in educational and psychological research.

6. **Factor Analysis**

Factor analysis is the term used to describe any one of a number of methods for analyzing the intercorrelations among a set of variables for the purpose of reducing the variables to as few dimensions as can be fruitfully used in order to describe the total set of variables under analysis. These techniques attempt to account for intercorrelations among variables in terms of underlying factors and to reveal the proportion of variation in each of the original measures associated with each hypothetical factor.

7. **Cluster Analysis**

Another data reduction technique, cluster analysis is a simplified version of factor analysis. As in factor analysis, the major focus of the technique is to account for intercorrelations among variables in terms of underlying factors. The major difference between cluster analysis and factor analysis is that, in cluster analysis, each variable as a unit is placed into a cluster, whereas in factor analysis, different portions of the variance of each variable may be assigned to different factors. As with factor analysis, there are a number of different cluster analysis techniques.

8. **Discriminant Analysis**

Discriminant analysis represents an extension of regression analysis to the case where the criterion variable is discrete rather than continuous. The model is generally used for classification purposes. In essence, it provides an estimate of the position of an individual, based on specified information on that individual, on a line that best separates two or more classes or groups. Since it is often the case that one best line may not exhaust the power of a given group of measures for discriminating among groups, additional discriminant functions (to the lesser of the number of groups minus one, or the number of measures) may be fitted. Thus, the major purpose of discriminant analysis is to determine whether discrimination among groups on the basis of a specified set of variables is possible or not, and then to reduce the size of the predictor space without substantial loss of information. Discriminant analysis itself does not define the regions of classification; however, approximate tests of the statistical significance of the separation of groups on a particular discriminant function are available, and the relative contributions of original variables to a discriminant function can be shown.

9. **Hotelling's T^2 and Mahalanobis' D^2**

These two multivariate tests are designed to investigate questions relative to the differences between group centroids. Hotelling's T^2, a generalization of Student's t-statistic to multivariate cases, and Mahalanobis' D^2, a measure of the distance between two group centroids, are both related to discriminant analysis for two groups. Both statistics may be used in situations where there are multiple measures on two groups, and yield identical probability statements. When the number of groups exceeds two, these tests are no longer applicable.

10. **Multivariate Analysis of Variance**

Multivariate analysis of variance is a statistical technique designed to analyze situations which involve both multiple independent and multiple dependent variables. Since the multivariate analysis of variance is the multivariate generalization of the univariate analysis of variance for testing the equality of mean vectors of several populations, design considerations are identical to those applied to the univariate analysis of variance. The difference is that the multivariate analysis of variance, a special case of the multivariate general linear hypothesis, obtains the probability of observed mean differences on more than one dependent variable simultaneously by an exact multivariate test of significance. Thus, a single probability statement applicable to all variables jointly is obtained.

Although it is true that separate univariate tests could be performed on each dependent variable, this procedure would generally not yield a single probability statement applicable to all variables jointly. Most dependent variables that are obtained from the same subjects will be correlated and, thus, will not yield statistically independent tests. No exact probability that at least one of the dependent variables will exceed some critical level on the null hypothesis can be calculated. Multivariate analysis of variance avoids this problem since the technique is based on sample statistics which take the correlations among dependent variables into account and have known sampling distributions (given that the assumptions are met) from which such probability statements can be obtained.

11. Canonical Analysis

Canonical analysis is a technique that can be used to study the interrelationships between two sets of measurements made on the same subjects. Both multiple criteria and predictors are involved. The technique maximizes the correlation between linear combinations of the two sets of variables. Canonical analysis is the multivariate generalization of univariate product-moment correlational analysis. Significance tests are available.

Due to the complexity of the task of finding two sets of combining weights, one for the predictor variables and a second for the criterion variables, the computational procedure is quite complex and yields one less solution than the lesser of the number of criterion or predictor variables.

12. Multivariate Regression Analysis

Multivariate regression analysis represents a generalization to the multivariate situation of univariate regression analysis. Thus, the technique is most useful in those situations where the investigation involves concomitant continuous variables that cannot be conveniently grouped into discrete categories without losing a great deal of information. The technique has many of the advantages of univariate regression analysis over classical analysis of variance, the most obvious being greater flexibility. In addition, procedures are available for determining whether or not the addition or deletion of given independent variables to the regression equation significantly improves prediction.

13. **Multivariate Analysis of Covariance**

Multivariate analysis of covariance represents a generalization to the multivariate situation of univariate analysis of covariance. As in the univariate case, multivariate analysis of covariance is not a method to adjust for lack of random sampling, but rather a technique that can be used to increase the precision of one's experiment. Design considerations are similar to those of univariate analysis of covariance.

14. **Kolmogorov-Smirnov One-Sample Test**

The Kolmogorov-Smirnov one-sample test is a non-parametric technique that is used to determine, within specified levels of probability, whether a set of observed scores can be considered to have been drawn from a population having some specified theoretical distribution. It is applicable for small or large samples and is more powerful than any alternative non-parametric test.

15. **One Sample Runs Test**

The one sample runs test is a non-parametric technique that is used to determine, within specified levels of probability, how likely it is that a sample represents a random picture of a population. The test focuses on the order of sample events and can be used for either small or large samples.

16. **Spearman Rank Order Correlation Coefficient (Rho)**

This statistic is a measure of association between ranks. It can be used for small or large samples, yields a coefficient between -1.00 and +1.00, and can be tested for statistical significance. It is an extremely powerful non-parametric technique for use with two ordinal variables.

17. **Kendall's Tau**

Kendall's Tau is a measure of association that is applicable to the same level of data as the Spearman rank order correlation coefficient, although numerically they are not comparable. However, when used in an inferential manner, both statistics have the same power to detect association in the population.

18. **Kendall's Partial Rank Correlation Coefficient**

This is superficially the non-parametric analog to the parametric product-moment partial correlation previously discussed (4). Essentially, this technique enables the researcher to statistically remove the effects of a third variable upon the relationship between two other variables. Since the distribution of the test statistic is unknown, it is not possible to test the computed measure of association for significance and, therefore, it cannot be said to have a parametric analog. Like Kendall's Tau, it is a measure of association for use with ranked data.

19. Sign Test

The sign test is applicable to situations in which there are two related samples and the evaluator wishes to determine, within specified levels of probability, whether the populations from which the samples were drawn share certain specified characteristics. The test focuses on the direction of differences between each pair of scores and is generally used to attempt to detect a shift in location. Probabilities are determined by reference to the binomial distribution. The power-efficiency of the sign test is high for small samples (n=6) but declines as sample size increases.

20. Wilcoxon Matched-Pairs Signed Ranks Test

This test is applicable to situations in which there are two related samples, and information relative to the magnitude as well as to the direction of differences is available. It is designed to determine, within specified levels of probability, whether the populations from which the samples were drawn share certain specified characteristics. This test requires ranked data and is an extremely powerful non-parametric technique.

21. Mann-Whitney U

This non-parametric technique has been developed independently by a great number of writers to test, within specified levels of probability, whether the populations from which two independent samples have been drawn have the same distribution. It is one of the most powerful non-parametric techniques and is a useful alternative to the parametric t-test.

22. Kolmogorov-Smirnov Two-Sample Test

The Kolmogorov-Smirnov two-sample test is a test of whether or not, within specified limits of probability, two independent samples have been drawn from populations with the same distribution. The test is sensitive to any kind of differences in the distributions from which the two samples were drawn and, thus, is concerned with the degree of isomorphism between two cumulative distributions. This test has a high power-efficiency for small samples, but it decreases slightly as sample size increases.

23. Wald-Wolfowitz Runs Test

This non-parametric technique is useful for determining, within specified limits of probability, whether the populations from which two independent samples have been drawn share certain characteristics against the alternative hypothesis that the two populations differ in any respect whatsoever. Although little is known about its power-efficiency, there is some evidence to suggest that it is not a very powerful technique.

24. **Kruskal-Wallis One-Way Analysis of Variance**

The Kruskal-Wallis one-way analysis of variance is useful for determining, within specified limits of probability, whether or not the populations from which two or more independent samples were drawn share certain specified characteristics. It is an extremely powerful non-parametric technique and a useful alternative to the parametric one-way analysis of variance. Procedures for post-hoc analysis have been developed.

25. **Kendall's Coefficient of Concordance**

The coefficient of concordance is useful for determining the association among a set of two or more rankings on two or more individuals. Such a measure is extremely useful for studies of interjudge reliability. The coefficient of concordance can be tested for significance.

26. **Friedman Two-Way Analysis of Variance**

The Friedman two-way analysis of variance is useful for determining, within specified levels of probabili y, whether the populations from which two or more related samples were drawn share specified characteristics. This is quite a powerful non-parametric test.

27. **The Binomial Test**

The binomial test is a goodness-of-fit test which determines, within specified limits of probability, whether it is reasonable to believe that the proportions that are observed in a given sample of only two classes could have been drawn from a population having some specified proportion of elements in each class. The binomial test assumes sampling with replacement (or sampling from an infinite population).

28. **The Hypergeometric Test**

The hypergeometric test is essentially designed to investigate the same questions as the binomial test but assumes sampling without replacement. Thus, it is more appropriate when sampling from a small finite population.

29. **Chi-Square One-Sample Test**

This test can handle two or more categories and is designed to determine, within specified levels of probability, whether a difference exists between an observed number of responses and an expected number, based on the null hypothesis.

30. **McNemar's Test for Change**

The McNemar's test for change determines, within specified levels of probability, whether or not observed changes in two related samples on the variable or variables of interest are greater than would be expected by chance.

31. Cochran Q Test

The Cochran Q Test represents an extension of the McNemar test for change to the situation where there are more than two related samples. It provides a statement, within specified limits of probability, of whether or not three or more matched sets of frequencies or proportions differ.

32. The Contingency Coefficient

The contingency coefficient is a measure of association that is closely related to Chi-square. It can be extended to situations where there are more than two samples and more than a dichotomization of the variable of interest and can be tested for significance.

33. Fisher's Exact Probability Test

This is a test for independent samples that determines, within specified limits of probability, how likely it is that two or more population proportions are equal. In theory, this test is appropriate, regardless of the sample size; in actuality, it becomes extremely tedious to compute with large sample sizes or with more than two samples.

34. Chi-Square

This is an extremely flexible and often used non-parametric statistical technique. It can be used to test the quality of population proportions (Chi-square test for homogeneity) or for statistical independence among related populations (Chi-square test for independence). Chi-square can be applied to more than two populations and more than a dichotomization of the variable or variables of interest. This is the only practical non-parametric technique in which it is possible to have both more than two samples and data which fall into more than two categories in the same design. Post-hoc tests are available for both Chi-square for homogeneity and for independence.

35. The Median Test

The median test is a non-parametric test which is designed to determine, within specified limits of probability, whether or not two independent samples have been drawn from populations with the same median. The median test uses Fisher's exact proability test for small samples and the Chi-square test for homogeneity for large samples, the median test itself merely being a method of categorizing data.

PROJECT DESCRIPTION

1. Describing the Context

A. Schools, as they have been organized traditionally, have emphasized abstract learning in fabricated situations. Our message to students has been that learning is something that happens inside schools. Rather than take advantage of adventures in reality we have found it easier to have students listen, or read, or look at pictures. This basic method of operation found in most schools creates a serious impediment when the subject is environmental education. The "ecological facts of life" must be woven into the fabric of every child's education, and education's responsibility begins as the child enters school.

Environmental education is rapidly becoming recognized as our nation's major curriculum concern. The intent of Project LOBO to increase the student's involvement with the natural environment in an organized manner and with some identifiable learning outcomes seems consistent with the major concerns of our society. Although the School City Schools were concerned about this topic, a coordinated effort was necessary to focus the work of the community and the schools relative to environmental education. The most appropriate source of funding for this creative endeavor was Title III of the Elementary and Secondary Education Act of 1965.

B. The purpose of this project is to broaden and enrich the base of activities in the School City community schools, both elementary and secondary, related to the understanding and preservation of the environment. The major emphasis is to develop a controlled system that provides for learning outside the classroom. Student awareness of the complexity of the environment should be increased through increased interaction with that environment.

C. School City is a lower middle class community located in central New Jersey whose population is primarily involved in vocational as opposed to professional and semi-professional occupations. Prior to the advent of Project LOBO the community per pupil expenditure was $970 for a total student population of 22,000.

2. Explaining the Project

A. Major Objectives:

(1) Students will engage in activities appropriate to their level of maturation which will include observation, investigation and evaluation of a variety of ecological relationships and conservation practices in central N.J. in order to develop the concept of stewardship of natural resources.

(2) Teachere will support the major objective and assist in its accomplishment as a result of the activities of this project.

B. Facilities Utilized:
(1) Two student transportation units each of which is a combination bus for about thirty students and a mini science lab. They serve as a link between the school and various sites in the greater community, and support and encourage a greater variety of field excursions of short duration as well as the work of the more sophisticated mobile laboratories.
(2) Two mobile laboratories for use at major sites utilized as important examples. These mobile units, forty-foot semi-trailers, include equipment and personnel which support on-site student investigations into various environments. Project LOBO also maintains office space in the administration offices of the School City Community Schools.

C. Specific Activities Include:
(1) Students investigate the micro-climate relative to temperature, water availability and type of vegetation as determined by insolation.
(2) While visiting a natural growth area and a cultivated area, students conduct soil investigations.
(3) Students investigate the earth's crust in order to determine the effects of former environments on various ecological relationships.
(4) Students investigate surface and subsurface water in a watershed that includes lakes, ponds and streams.
(5) While visiting a timber tract, students investigate the complexity of the interrelationships of life in a forest.
(6) Students investigate the economic mineral resources of gravel, limestone, ag lime and gypsum while on site.
(7) Students investigate methods used in plant-animal conservation.
(8) Students investigate methods used in soil conservation.
(9) Students investigate methods used in water conservation.

D. During the 1981-82 school year, 3,181 students and 121 teachers participated in project field activities. This resulted in 43,548 student hours in the field. These students were from grades 1,3,5,7 and 9, and their classroom teachers each spent twelve hours in in-service activities.

The teacher in-service program is coordinated with the seasons and with the three trips by their students. Prior to a class level program beginning, the teachers of that level are released from duties and spend one-half day in instruction specific to the location that the class will be visiting and activities appropriate for that location. This not only assists the teachers in the preparation of the class for the experiences but also gives them confidence that they can operate in a setting different than the classroom.

The second portion of the in-service program is more subtle but equally important. The field trip itself with class and teacher provides an opportunity for observing the project

staff member as a model, not just for the techniques of presenting information, but as a master teacher in such areas as the relationship established with students and creation of an open, discovering environment for learning. Thus the twelve hours a year in direct "instruction" in the field is completed with another eighteen hours of modeling behavior while the teacher is in the field with her class. Teachers have been generous in their praise for this form of in-service program.

E. Human Interest:

There are several areas of project endeavors where we have been pleasantly surprised with acticities that have exceeded our original expectations. An all-encompassing area is that of people-reactions. We firmly believed that students would be "turned on" to our field activities and reactions have been on target. We did not anticipate the extent of teacher appreciation that has occirred, nor did we anticipate the several dramatic changes that have occurred in the lives of some students.

Teachers have reported to us that they now have more confidence in themselves and their ability to function with a class in the field as they provide significant learning experiences for the students. They also have commented most favorably on the effect that the project had on the large segment of students who could be described as quiet, almost withdrawn; slow learners; lacking high personal motivation. These students who apparently were not functioning well in the classroom reversed their position on the field trips. This newly-found status was carried back to their classrooms to the benefit of the entire group.

A specific example is the story of one nine-year old boy who early in the year was considered to have such serious problems that his continuation in the public school was questioned. He was so successful as a good worker on the LOBO field trips that he was recognized by the project staff with a certificate of merit. The teacher contends that his involvement in the project has given the boy a basis for working out his other problems, and he is making excellent progress in his class.

Another example is the story of a high school boy who, as a result of contacts made through the School City High Ecology Club, has volunteered his assistance three afternoons a week. He has helped with students in the field, with the organization of the central store of equipment and helped organize the summer activities as well as other ecological activities of his school. Unsolicited reports from both his family and his school counselor indicate that his involvement in the project has improved his self-concept and given him a purpose for his school activities. These examples, though dramatic, are just two of the several that might be reported.

Cooperation with Project LOBO in making sites available for field study has been excellent. The Morton County Conservation Board has been most cooperative, and in a letter to the project staff they expressed their appreciation of the project's efforts to properly use the county parks as outdoor classrooms by stating, "It is through such programs as yours that parks are brought to the people, and a fuller appreciation thereof results."

3. Describing Effectiveness

A. Strategy:

The major objective is seen as a global goal which has been refined to student behaviors through two main channels--one cognitive and one affective. Students need knowledge of the total environment in order to build an attitude of "stewardship of natural resources." Assessment activities followed these two channels with both students and teachers. Parental attitude, necessary for further support, was also assessed.

Students and teachers in what have traditionally been called grades 1,3,5,7 and 9 were the targets for the project program during the first year of pperation. Control and experimental groups were randomly assigned by lottery in August, 1980. At the elementary level, five schools were chosen for study at each of the three levels. A random selection of classes utilizing the following guidelines was then made: 2 of 4 one-unit schools; 2 of 4 two-unit schools; and one of 2 three-unit schools. A third three-unit school was not included in this research design due to the involvement as a specific component of Project LOBO. At the junior high level, all the seventh and ninth grade science teachers were included although only one-half of the classes that these individuals teach were included in the treatment group.

B. Evaluation Results:

"Success is a process involving the progressive realization of worthwhile goals." If this definition, taken from a speech by Mr. Greg Woznick of Success Motivation Institute, Inc., is acceptable, then the first year of operation of Project LOBO was a success. The project is making progress toward the identified goals, and objective testing indicates that students and teachers in the system are making gains in areas of knowledge and attitudes.

These results are contained in a separate research report published by the project staff, but a few brief statements can be made. Testing with students at all levels revealed improved attitudes in the experimental groups relative to environmental management education. Improvements in cognitive test results at best favored the experimental groups and at worst revealed no significant differences between groups. Teacher attitudes and knowledge were improved as a result of the project activities. Parent attitudes were very supportive of Project LOBO activities.

Specific information on evaluation results at each level can be outlined as follows:
(1) Primary students: At the Primary II level (first grade), the control group scored significantly higher than the experimental group on the pre-test, indicating that they were "beginning" the year ahead of the experimental group. The fact that the post-test scores of the two groups were almost identical and that both groups recorded significant gains indicates that something important occurred in the learning of the students in the experimental group.

At the Primary IV level (third grade), the experimental group scored higher than the control group on the pre-test and the magnitude of their gain was greater than the control group. The gain scores of both groups were significant, but the difference between groups was not significant. What of the gain scores of primary students was due to maturity and what due to project contamination is not known. Although this does not allow neat research-based postulations, it is valued by those concerned with the end product of the total system.

(2) Intermediate students: The test item development procedure undertaken with this group has not yet revealed objective data. The process has been valuable to teachers and to students to the extent that both groups have been involved in sharing information about what was "learned" as a result of project activities. A very substantial result of this process is a computer-based bank of 150 questions which can be "summoned out" by teachers. The Stanford Achievement Test results as revealed in a comparison of the building level mean scores on the science subtest in 1981 seem to favor the experimental group. Four of the five experimental groups scored a higher mean score than their control group counterparts, and the one experimental school whose mean score was lower had narrowed the gap that existed in previous years.

(3) Junior high students: Results of assessment in the cognitive and affective domains were favorable for the groups of students included in Project LOBO activities (Experimental). There were no significant differences in tests of knowledge, either pre or post, for the control and experimental groups at grades seven and nine, although the internal gains for each group were significant and the experimental to a greater magnitude. The environmental survey revealed a significantly less desirable attitude on the part of the seventh grade experimental group on the pre-activity assessment and a significant improvement in their mean scores in relation to the control group, which regressed. At the ninth grade level, significant improvements occurred in attitude scores for both control and experimental groups, and again the difference, though not significant, favored the experimental groups.

(4) Teacher in-service: The knowledge level and attitude toward environmental education evidenced improvement on the part of School City teachers. Teachers in the experimental group were unanimous in their approval of Project LOBO as a "worthwhile curriculum venture." The request to participate the second year on the part of all teachers is further evidence of the value placement of the project.

The construction of an attitude instrument for teachers utilized the work of Roth, et al,* to identify a list of environmental management concepts appropriate to Project LOBO objectives. The 67 concepts identified were put in a form that asked respondents how important they believed the concept to be for their teaching level and how adequate they felt relative to trying to teach that concept. The use of the Cronbach-Alpha Formula indicated reliability coefficient figures ranging from .94 to .97 among the groups to whom the instrument was administered.

(5) Parent attitudes: A questionnaire was sent to the parents of participants at the end of the first year of Project LOBO and at the end of the second year of operation. We now have over 1,100 responses which contain evidence of the community being favorably impressed and supportive of the project. Over 90% agree field trips are important; that LOBO provides worthwhile experiences; that LOBO did NOT detract from more worthwhile classroom activities; that their children talked about the trips; that the activities were enjoyed; that the curriculum venture is worthwhile; and that they want their children included next year.

The response to the question as to whether the community would support LOBO without Federal funding is quite positive:
(a) At the conclusion of the first year 16% disagreed with the statement that the community would support the project if it was not Federally funded, and 34% agreed or strongly agreed.
(b) At the conclusion of the second year 18% of the elementary parents disagreed, but the agreement responses had reached 40%.
(c) At the conclusion of the second year 22% of the junior high parents disagreed, but the agreement responses had reached 46%.

The project has benefited many curriculum areas. In a mid-year teacher survey to thirty representative teachers from the elementary adn junior high classes the following results were obtained:
(a) Eighty percent felt that there was carry-over in language arts.

*Roth, Robert E., ...

(b) Sixty-six percent felt there was carry-over in social studies and the science areas.
(c) Sixty percent felt there was carry-over in math and art.
Teachers also mentioned carry-over in the areas of emotional health and physical education.

4. Describing Costs

A. Developmental costs are estimated to be $6,500. This is approximately one-half of the expenses of our first year and seems appropriate if our experiences are utilized, since much of the initial expense was related to proposal writing and production.

B. Initiation or start-up costs are estimated at $60,000-$80,000. This variance is related to the amount of utilization of existing district equipment and transportation facilities. The School City Community Schools had no buses, so the entire transportation portion of the project was included in start-up costs.

C. Operational costs after installation are estimated at $48,000. Salaries of professional and non-professional staff and staff development (new teacher orientation) would constitute spproximately $42,000 of this figure. The $6,000 will take care of equipment maintenance and transportation.

5. Describing Exportability Factors

Very few special factors need to be considered when an LEA is planning to adopt a project such as LOBO. Certainly the project staff is important, but any school district would have interested and committed individuals who would be knowledgeable enough to provide leadership in this area.

Outdoor sites must be available, but practically all communities can provide necessary parks, lakes and representative ecological areas as well as examples of man's intervention in the environment.

Project LOBO was fortunate to have developed in a school district that had a talented maintenance staff that was capable of remodeling the used forty-foot semi-trailers so that this was not a prohibitive expense.

Consideration should also be given to the amount and level of environmental education that is going on in the school system at the time. This gives some index of the readiness of the staff and the community for increased activities.

Although is is the Project LOBO staff's belief that the costs associated with such a project are moderate, there is an awareness of the budget limitations placed on local schools by legislatures in many states. Adjustments can be made which

would reflect the priority that a school system places on environmental education. The emphasis of many curriculum areas, at both the elementary and the secondary level, can be directed toward environmental "awareness" with little added expense. Field trips are an integral part of most schools' activities, and the shift of emphasis again is not overly expensive. The provision of support-lab facilities and some monitoring equipment for field investigations is a special consideration that needs to be faced by anyone wanting to "import" Project LOBO.

6. Publications and Materials

The project staff publishes a monthly newsletter which is currently used as a "handout" to describe the project in general. A brochure has been produced that presents a brief view of the major elements of developing a forty-foot semi-trailer into a mobile science laboratory for field work. This goes from purchase through remodeling of the interior and the final equipping of the unit. Also available at a cost of $10.00 is a spiral bound Teacher Resource Booklet. This 145-page manual contains specific material organized separately for grades 1,3,5,7 and 9, but applicable to a broader age span, relating to activities and ideas for projects to be conducted in the field. Slide and tape presentations, illustrating Project LOBO activities, are in the process of development.

7. Describing Unanticipated Outcomes and Spinoff Findings

An encouraging and rewarding outcome of the project is the National Science Foundation funding of a Cooperative College-School Science Program developed by staff members from Fink University and members of the LOBO staff. This project has enabled 25 elementary school teachers from the School City Community School District to be participants in a three-week teacher training program during the summer of 1982. The program is designed to train a nucleus of teachers in the School City school system who will develop a working knowledge of Project LOBO and provide in-service training for their peers. During the 1982-83 academic year these participants will meet monthly for further instruction and group interaction. The development of this project further emphasizes the excellent cooperation the LOBO staff has enjoyed with staff members of Fink University.

Validator Self-Analysis Form

Please fill in the following information at the conclusion of the on-site validation and mail in the envelope enclosed with the Guidebook.

Name_____ Date_____

Project Reviewed for Validation_____

City_____ State_____

Your Address_____

City_____ State_____

Section Reviewed_____

1. Were you a validator last year? Yes_____ No____

2. If yes, did you validate the same criterion as last year? Yes_____ No____

3. Do you feel your involvement reflected your area of expertise? Yes_____ No____

4. Were the point values for the section you validated generally acceptable to your teammates? Yes_____ No____

5. Did you find the task of assessing the data and weighing the responses difficult? Yes_____ No____

   Comments:_____

   _____

   _____

6. Is this year's validation instrument generally better or worse than last year's? Better____ Worse____

7. Which section of the instrument was most improved over last year's?

   Effectiveness/Success_____ Cost Information_____ Exportability____

8. Were there questions in your section that you found particularly difficult to answer in terms of assessing number weights? Yes___ No___
   If yes, please indicate section and question numbers

   _____

   _____

9. Do you feel that there was adequate team interaction and discussion in reaching a conclusion on each of the three sections of the report?
   Yes_____ No____

10. Please comment on areas of difficulty with respect to both the validation instrument, and the team's interaction with one another and with project personnel.

    _____

    _____

STATE CODES

ALABAMA	01
ALASKA	02
ARIZONA	03
ARKANSAS	04
CALIFORNIA	05
COLORADO	06
CONNECTICUT	07
DELAWARE	08
DISTRICT OF COLUMBIA	09
FLORIDA	10
GEORGIA	11
HAWAII	12
IDAHO	13
ILLINOIS	14
INDIANA	15
IOWA	16
KANSAS	17
KENTUCKY	18
LOUISIANA	19
MAINE	20
MARYLAND	21
MASSACHUSETTS	22
MICHIGAN	23
MINNESOTA	24
MISSISSIPPI	25
MISSOURI	26
MONTANA	27
NEBRASKA	28
NEVADA	29
NEW HAMPSHIRE	30
NEW JERSEY	31
NEW MEXICO	32
NEW YORK	33
NORTH CAROLINA	34
NORTH DAKOTA	35
OHIO	36
OKLAHOMA	37
OREGON	38
PENNSYLVANIA	39
RHODE ISLAND	40
SOUTH CAROLINA	41
SOUTH DAKOTA	42
TENNESSEE	43
TEXAS	44
UTAH	45
VERMONT	46
VIRGINIA	47
WASHINGTON	48
WEST VIRGINIA	49
WISCONSIN	50
WYOMING	51
BUREAU OF INDIAN AFFAIRS	52
AMERICAN SAMOA	53
VIRGIN ISLANDS	54
GUAM	55
PUERTO RICO	56
TRUST TERRITORY	57

BIBLIOGRAPHY

Please note: The books included on this list represent a cross section of materials considered helpful in the conduct of educational evaluation. The titles indicated by an asterisk (*) have been utilized in either the Guidebook or the training materials.

E. F. Shelley & Co. Inc. "An Educational Audit System" 415 Madison Ave., NYC - 1970

Cronback, Lee J., "Essentials of Psychological Testing" Harper & Row, New York - 1970

"Elementary School Test Evaluations" Center for the Study of Evaluation, UCLA Graduate School of Education, Los Angeles, Cal. - 1970

*Glass, G., and Stanley, J., "Statistical Methods in Education and Psychology" Prentice Hall, Englewood Cliffs, N.J. - 1970

Kerlinger, F.N., "Foundations of Behavioral Research" Holt, Rinehart & Winston, New York - 1964

Provus, Malcolm "Discrepancy Evaluation" Mc Cutcheon Press, Berkeley, Cal. - 1971

Seltiz, C., Jahoda, M., Deutsch, M., & Cook, S.W. "Research Methods in Social Relations" Holt, Rinehart & Winston, New York - 1959

Stenner, A.J. and Webster, W.J. "Educational Program Audit Handbook" Institute for the Development of Educational Auditing, 1121 Arlington Blvd., Arlington, Va. 22209 - 1971

*Stufflebeam,Daniel, et al. "Educational Evaluation and Decision-Making" Peacock Publishers, Inc. Itasca, Ill. - 1971

"Tests & Indicators: Ways to Find Out Where You Are", N.J. Dept. of Education: Office of Program Development, Trenton, N.J. - 1972

Title VIII, ESEA Manual for Project Applicants & Grantees USOE, Washington, D.C. 20202 (section on responsibilities of the educational auditor)

Tyler, Ralph W. (ed.) "Educational Evaluation: New Roles, New Means" University of Chicago Press, Chicago, ILL.- 1969

Weiss, C.H., "Evaluating Action Programs" Allyn and Bacon, Boston, Mass. - 1972

* Havelock, R.G., "Planning For Innovation Through The Dissemination and Utilization of Knowledge" Univ. of Michigan Press, Ann Arbor, Michigan - 1969

APPENDIX C

Note:

The following list of projects were candidates for Validation, 1973-74. It was compiled from the completed and returned Self-Analysis Forms and by definition incomplete. The list does not intend to convey that any one or all projects appearing on the list were validated.

State	City	Name of Project
Alabama	Mobile	Talents Unlimited
Alaska	Atarhak	Early Childhood Project
Arkansas	Rogers	Rogers Parent Education Research Center
California	San Diego	Multigrade Grouping in Early Childhood Education
Colorado	Colorado Springs	Parade (Reading and Learning)
Colorado	Colorado Springs	Interdisciplinary Career Education
Colorado	Lakewood	Added Dimensions to Preschool & Parent Education
Florida	St. Petersburg	Pupil Personnel Services
Georgia	Columbus	Learning Music as a Language
Georgia	Scottdale	Program for Child with Specific Learning Disabilities
Georgia	Waycross	Pride
Georgia	Waycross	Serving Young Children with Cross Disabilities
Illinois	Belleville	Project "Reach"
Illinois	Quincy	Understanding: An Environmental Structure for Education (PIE)
Illinois	Urbana	Junior High School Reading Laboratory
Indiana	Hammond	Unifon Reading
Indiana	Marion	Project Puntoon III
Iowa	Cedar Rapids	Basic Reading and Staff Development
Iowa	Burlington	Clinical Speech Services
Iowa	Des Moines	A Diagnostic and Educational Center for Learning Problems
Kentucky	Morganfield	Total Phased Curriculum
Kentucky	Nashville	A Prevention-Intervention Model for Students' Learning and Behavior Problems

State	City	Name of Project
Kansas	Turner (Kansas City)	Individualized Instruction in Family Living
Kansas	Pittsburg	S.E.T. Project
Kansas	Wichita	Project Deep
Kansas	St. John	Education for the High Performance Pupil
Kansas	Topeka	Environmental Education Demonstration Prod.
Louisiana	Baton Rouge	Talents Unlimited
Louisiana	Lafayette	Program for Low Achievers in Math
Maryland	Rockville	Early Childhood Services for Visually Impaired Children
Michigan	Berrien Springs	IMPACT: Instructional Model for all Children and Teachers
Michigan	Marysville	Haptic Perceptual Development
Minnesota	Mora	Mora 45-15 Elementary School Schedule
Missouri	Maplewood	Maplewood-Richmond Heights Pre-School Program
Missouri	Cape Girardeau	Facilitating Learning through Systems Modification
Nebraska	Bellevue	Model Guidance
Nebraska	Bellevue	Right to Read
Nebraska	Holdrege	Project W.O.R.C.
Nebraska	Arnold	Video Taped Packages - Math
Nebraska	Omaha	Empathy
Nebraska	Alliance	Indian Guidance Project
Nebraska	St. Edward	St. Edward Preschool
New Hampshire	Gorham	North Country Education Services
New Jersey	Pitman	Institute for Political/Legal Education
New Jersey	West Long Branch	Project ACTIVE

State	City	Name of Project
New York	Guilderland	Instructional Support System
North Carolina	Albemarle	Project LAD (Learning Abilities Development)
North Dakota	Lignite	Northwest Special Education Program Model
North Dakota	Minot	Early Identification of Learner Need
North Dakota	Fargo	Pre-Kindergarten Prescriptive Teaching
Ohio	Akron	Robinson Title III Project
Oklahoma	Edmond	Accountability and Mini-Course
Oklahoma	Sapulpa	Success Through Identification and Curriculum Change
Oregon	Vale	Occupational Education for Non-College Bound Students
Pennsylvania	West Chester	Cognitively Oriented Urban Pre-Kindergarten
South Carolina	Florence	Individualized Instruction
Texas	Rockwell	Social Problems of Today
Texas	Richardson	The MOD Project
Washington	Yakima	Studio Study Center for Creatively Talented Students
Washington	Poulsbo	Project Success
Washington	Everett	Project Turnabout
Washington	Port Townsend	Cooperative for Handicapped Students
Wisconsin	Waupun	Comparative Cultures
Wisconsin	Green Bay	Instruction--Curriculum--Environment